MAYBE

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

FILE

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 9/22, C07K 14/195, C12N 15/55, C12Q 1/68	A1	(11) International Publication Number: WO 00/17330
		(43) International Publication Date: 30 March 2000 (30.03.00)

(54) Title: DNA POLYMERASE HAVING ABILITY TO REDUCE INNATE SELECTIVE DISCRIMINATION AGAINST FLUORESCENT DYE–LABELED DIDEOXYNUCLEOTIDES

(57) Abstract

The invention relates to genetical modification of DNA polymerase to reduce its innate selective sequence–related discrimination against incorporation of fluorescent dye–labeled ddCTP and ddATP in the enzymatic reaction for preparation of samples for automated fluorescent dye–labeled terminator DNA sequencing. The modified DNA polymerases are more resistant to heat inactivation and are more effective in dideoxynucleotide incorporation than current DNA polymerases.

DNA POLYMERASE HAVING ABILITY TO REDUCE INNATE
SELECTIVE DISCRIMINATION AGAINST
FLUORESCENT DYE-LABELED DIDEOXYNUCLEOTIDES

BACKGROUND OF THE INVENTION

The genetic material of all known living
organisms is deoxyribonucleic acid (DNA), except
in certain viruses whose genetic material may be
ribonucleic acid (RNA). DNA consists of a chain of
5   individual deoxynucleotides chemically linked in
specific sequences. Each deoxynucleotide contains
one of the four nitrogenous bases which may be
adenine (A), cytosine (C), guanine (G) or thymine
(T), and a deoxyribose, which is a pentose, with a
10   hydroxyl group attached to its 3' position and a
phosphate group attached to its 5' position. The
contiguous deoxynucleotides that form the DNA
chain are connected to each other by a
phosphodiester bond linking the 5' position of one
15   pentose ring to the 3' position of the next
pentose ring in such a manner that the beginning
of the DNA molecule always has a phosphate group
attached to the 5' carbon of a deoxyribose. The
end of the DNA molecule always has an OH
20   (hydroxyl) group on the 3' carbon of a
deoxyribose.

DNA usually exists as a double-stranded
molecule in which two antiparallel DNA strands are
held together by hydrogen bonds between the bases
25   of the individual nucleotides of the two DNA
strands in a strictly matched "A-T " and "C-G"
pairing manner. It is the order or sequence of the
bases in a strand of DNA that determines a gene
which in turn determines the type of protein to be
30   synthesized. Therefore, the accurate determination
of the sequence of the bases in a DNA strand which

also constitutes the genetic code for a protein is
of fundamental importance in understanding the
characteristics of the protein concerned.

5   The process used to determine the
sequence of the bases in a DNA molecule is
referred to as DNA sequencing. Among the
techniques of DNA sequencing, the enzymatic method
developed by Sanger et al. (1) is most popular. It
is based on the ability of a DNA polymerase to
10   extend a primer annealed to the DNA template to be
sequenced in the presence of four normal
deoxynucleotide triphosphates (dNTPs), namely,
dATP, dCTP, dGTP and dTTP, and on the ability of
the nucleotide analogs, the dideoxynucleotide
15   triphosphates (ddNTPs), namely, ddATP, ddCTP,
ddGTP and ddTTP, to terminate the extension of the
elongating deoxynucleotide polymers at various
lengths.

In the classic one-step Sanger method,
20   the sequence determination is carried out in a set
of four separate tubes, each containing all four
normal dNTPs, one of which is labeled with a
radioactive isotope, ^{32}P or ^{35}S, for
autoradiographic localization, a limiting amount
25   of one of the four ddNTPs, a DNA polymerase, a
primer, and the DNA template to be sequenced. As a
result of the DNA polymerase activity, individual
nucleotides or nucleotide analogs are added to the
new DNA chains, all starting from the 3' end of
30   the primer in a 5'-3' direction, and each linked
to adjacent ones with a phosphodiester bond in a
base sequence complementary to the DNA sequence of
the template. Inasmuch as there is a nucleotide
analog in the reaction mixture, each tube
35   eventually contains numerous newly formed DNA
strands of various lengths, all ending in a
particular ddNTP, referred to as A, C, G or T

2

terminator.

After resolving the four sets of
reaction products by high-resolution
polyacrylamide/urea gel electrophoresis, the
5   populations of the newly formed DNA strands are
separated and grouped according to their molecular
weight. An autoradiographic image of the gel will
show the relative positions of these DNA strands
as bands which differ from one another in distance
10   measured by one nucleotide in length, all sharing
an identical primer and terminating with a
particular ddNTP (A, C ,G or T). By reading the
relative positions of these bands in the "ladder"
of the autoradiograph, the DNA sequence of the
15   template can be deduced.

The DNA polymerase used in the reaction
mixture plays a pivotal role in DNA sequencing
analysis. To be useful for DNA sequencing, a DNA
polymerase must possess certain essential
20   properties. For example, it must have its natural
5'-3' exonuclease activity removed by mutagenesis
or by posttranslational modification, such as
enzymatic digestion, and must be able to
incorporate dNTPs and ddNTPs, without undue
25   discrimination against ddNTP and with a
sufficiently high processivity which refers to the
ability of the enzyme to polymerize nucleotides
onto a DNA chain continuously without being
dislodged from the chain, and a sufficiently high
30   elongation rate. A 5'-3' exonuclease activity
associated with a DNA polymerase will remove
nucleotides from the primer, thus cause a
heterogeneous 5' end for the newly formed DNA
strands, resulting in a false reading of the
35   strand lengths on the sequencing gel. A DNA
polymerase with a low processivity and a low
elongation rate will cause many undesirable noise

3

background bands of radioactivity due to the
presence of DNA strands which are formed with
improper lengths and improper terminations. Among
the more commonly used DNA polymerases, Sequenase™

5   has a higher processivity and a higher elongation
rate than others, such as the Klenow fragment,
Taq, and Vent polymerases (2), and is therefore
one of the most popular DNA polymerase selected
for DNA sequencing to-date.

10   However, even when a DNA polymerase has
been endowed with all the essential properties
listed above, it may still generate erroneous or
misleading band patterns of radioactivity in the
sequencing gel. These artifactual patterns do not

15   faithfully reflect the true nucleotide sequence in
the template being sequenced. They may be caused
by premature termination of the elongating strands
due to the presence of secondary structures formed
along the template, such as " hairpins " in the

20   regions that contain palindromic sequences or that
are rich in G and C bases (3); or, they may occur
as a result of inadequate " proof-reading "
function of the DNA polymerase that will allow the
removal of misincorporated nucleotides at the 3'

25   end of an elongating strand.
             Researchers in the field of DNA
sequencing often have to use several approaches to
confirm their findings in order to avoid being
misled by these potentially erroneous sequence

30   data. For example, they sometimes rely on
repeating the same sequencing experiment with
different DNA polymerases, or performing another
sequencing reaction with the template which is
complementary to the first single-stranded DNA

35   template, and compare the results for possible
discrepancies.
             Numerous investigators have tried to

4

find an ideal DNA polymerase for enzymatic
sequencing, i.e. an enzyme that not only has all
the essential properties required for sequencing
reaction, but also is capable of resolving the
5   secondary hairpin structures and preventing the
formation of strands containing nucleotides non-
complementary to those of the template being
sequenced.

The discovery by Ye and Hong (4) of the
10   thermostable large fragment of DNA polymerase
isolated from *Bacillus stearothermophilus* (Bst),
an enzyme that is functional over the temperature
range between 25°C and 75°C, but is most active at
65°C, and possesses all the essential properties
15   for DNA sequencing, has largely solved the problem
caused by secondary structures in the template
since these secondary structures are destabilized
when the sequencing reaction is carried out at
65°C. In the past few years since this enzyme was
20   made commercially available under the name of Bst
DNA Polymerase (Bio-Rad Laboratories), independent
reports have confirmed that during sequencing
reaction catalyzed by this enzyme all four dNTPs,
including dCTP, and other nucleotide analogs, such
25   as dITP and 7-deaza-dGTP, are incorporated equally
effectively in the chain elongation, thus
eliminating the weak "C" band phenomena often
observed when other DNA polymerases are used, and
producing a very good band uniformity on the
30   sequencing gel. It has been further established
that at this elevated temperature Bst DNA
polymerase system can be used both for the classic
Sanger one-step reaction as well as for the
"labeling/termination" sequencing reaction,
35   double-stranded DNA sequencing, and the
incorporation of ^{35}S-labeled nucleotides, and ^{32}P-
labeled nucleotides. Since this system can be

placed at room temperature for at least two weeks
without significant loss of its enzymatic
activity, it has been adapted for automation of
DNA sequencing which requires a stable DNA

5   polymerase, using either fluorescent dye or
radioactive isotope labeling. (See also 9, 12, and
13.)

However, when this Bst enzyme is used
for automated fluorescent DNA sequencing, only

10   partially satisfactory results have been obtained
with fluorescent dye-labeled primers (see 12 and
EG Bulletin 1771 of Bio-Rad Laboratories), and
even less satisfactory results are obtained with
fluorescent dye-labeled ddNTP terminators. Even

15   when fluorescent dye-labeled primers are used, a
significant number of mismatched ddNTPs are
incorporated onto the 3' end of the extending
nucleotides in the enzymatic reaction, thus
generating erroneous sequencing data (see Bio-Rad

20   EG Bulletin 1771). With this in mind, the
inventors sought, and found, a better DNA
polymerase for DNA sequencing, especially for
automated fluorescent dye-labeled primer and
fluorescent dye-labeled terminator sequencing.

25   Another disadvantage of the Bst DNA
polymerase currently known in the art is its lack
of 3'-5' exonuclease activity (5), and
specifically, proof-reading 3'-5' exonuclease
activity. A survey of the sequencing data

30   collected from fourteen research centers which
have used this Bst DNA polymerase for their DNA
sequencing work on over 120 DNA clones showed
that, statistically, base pair mismatching occurs
at a rate of about 1.5×10^{-5}. That is,

35   approximately 1.5 errors can be expected in one
hundred thousand nucleotide incorporations during
nucleotide polymerization catalyzed by the enzyme.

It is generally known that the formation
of incorrect DNA sequences due to mismatching of
base pairs between the template and the growing
nucleotide chain in DNA sequencing may be
5   prevented by a 3'-5' exonuclease activity which
"proof-reads" the nucleotide chain. However, even
if a DNA polymerase exhibits 3'-5' exonuclease
activity *in vitro*, it is often the case that the
polymerase will not adequately "proof-read".
10   Thus, the polymerase will not be capable of
removing mismatched nucleotides from a newly
formed DNA strand as efficiently as those
nucleotides correctly matched with the nucleotides
of the template. In other words, a 3'-5'
15   exonuclease may excise the correctly matched
nucleotides at a faster rate than the mismatched
ones from the 3' terminus, or excise both the
correctly matched and the mismatched nucleotides
at the same rate. Consequently, even where the
20   DNA polymerase has 3'-5' exonuclease activity, it
does not perform any useful proof-reading function
during DNA polymerization.

It is also known that a 3'-5'
exonuclease activity associated with a DNA
25   polymerase, in the presence of low concentrations
of dNTPs, often counteracts the normal chain
elongation process catalyzed by the polymerase,
induces cyclic incorporation and degradation of
nucleotides over the same segment of template, or
30   even operates more efficiently than the polymerase
activity *per se*, to the extent of causing
degradation of the primer. Consequently, removal
of the 3'-5' exonuclease activity along with the
5'-3' exonuclease activity from the native DNA
35   polymerases by chemical means or by genetic
engineering techniques has become a standard
procedure in producing DNA polymerases for

7

sequencing. This is a common strategy to preserve the essential properties of a DNA polymerase.

For example, among the major commercially available sequencing enzymes (other than the native Taq (*Thermus aquaticus*) DNA polymerase which lacks a 3'-5' exonuclease activity *de novo*) the 3'-5' exonuclease activity has been removed from the native T7 DNA polymerase, which lacks a 5'-3' exonuclease, either by a chemical reaction that oxidizes the amino acid residues essential for the exonuclease activity (Sequenase™ Version 1) or genetically by deleting 28 amino acids essential for the 3'-5' exonuclease activity (Sequenase™ 2).

Vent$_R$(exo⁻) DNA polymerase, which is recommended as the preferred form of the Vent DNA polymerase for sequencing, also has its 3'-5' exonuclease activity removed by genetic modification. The native Vent DNA polymerase and the Klenow fragment isolated from the native E. coli DNA polymerase I possess a 3'-5' exonuclease; but these enzymes are no longer considered the enzymes of choice for DNA sequencing.

The currently known Bst DNA polymerase (e.g., produced by Bio-Rad Laboratories) isolated and purified from the cells of *Bacillus stearothermophilus* for DNA sequencing is free of 3'-5' exonuclease activity (5).

IsoTherm™ DNA Polymerase, a commercially available Bst DNA polymerase for DNA sequencing, marketed by Epicentre Technologies (1402 Emil Street, Madison, WI 53713), is also based on a Bst DNA polymerase whose 3'-5' exonuclease activity has been enzymatically removed (6).

Only the rBst DNA Polymerase produced from an over-expressing recombinant clone in E. coli, which is the product of the DNA pol I gene

8

of *Bacillus stearothermophilus*, possesses a 3'-5'
exonuclease activity in addition to a 5'-3'
exonuclease activity. However, due to the
existence of an undesirable 5'-3' exonuclease

5   activity and a 3'-5' exonuclease activity of
unknown characteristics, the latter product is not
recommended by the company for DNA sequencing (6).

Over the past 10 years there has been a
trend to develop and improve the automated

10 fluorescent DNA sequencing technology to replace
the classic radioactive isotope labeling manual
method for DNA sequencing because of the potential
harmful effects of the radioactive materials to
humans and because of the need for automated high

15 throughput DNA sequencing systems. In using
fluorescent dyes as markers for labeling the DNA
strands generated in enzymatic reactions for
sequencing, the dyes can be either coupled with
the primer, or coupled with the ddNTP terminators,

20 namely the dye-labeled ddATP, dye-labeled ddCTP,
dye-labeled ddGTP and dye-labeled ddTTP.
Sequencing techniques based on these two forms of
labeling of the final enzymatic reaction products
are commonly referred to as "dye primer

25 sequencing" and "dye terminator sequencing",
respectively.

In the dye primer sequencing, ddNTPs are
employed as the chain terminators, as in the
original classic Sanger method which uses

30 radioactive isotope as the marker. The molecular
structure of ddNTPs are almost identical to that
of dNTPs, the natural building blocks of all DNA
molecules. Therefore, any DNA polymerase which
has been used for radioactive isotope manual DNA

35 sequencing can be easily adapted for fluorescent
dye primer DNA sequencing with equally
satisfactory results. The disadvantage in the dye

9

primer technology is that the primer for each
template to be sequenced must be labeled with four
different fluorescent dyes and that the enzymatic
reaction must be performed in four separate test

5   tubes each containing only one of the ddNTPs,
namely ddATP, ddCTP, ddGTP or ddTTP, as in the
classic Sanger radioisotope method.

In the dye terminator technology for DNA
sequencing, the fluorescent dye-labeled ddATP,

10   dye-labeled ddCTP, dye-labeled ddGTP and dye-
labeled ddTTP are coupled with different
fluorescent dyes, each emitting a specific light
spectrum, thus directly reporting the type of
ddNTP at the 3' terminus of the DNA fragment.

15   Unlike the situations in the dye primer technology
in which four different fluorescent dyes are
coupled to a primer incorporated into all newly
formed DNA strands, these dye-labeled ddNTPs serve
the dual function of a specific base terminator

20   and a "color marker". There is no need to label
the primer for each new template, and the
polymerase DNA extension reaction can be performed
in a single test tube to generate the required
specifically terminated and specifically dye-

25   labeled DNA fragments of various sizes for DNA
sequencing.

The advantage of using fluorescent dye-
labeled terminators for DNA sequencing is obvious.
However, there are certain difficulties to

30   overcome before an enzymatic reaction system
suitable for a radioisotope technique or suitable
for a dye primer technique can be adapted for a
dye terminator technology. An increase of the
molecular weight from less than 500 for a ddNTP

35   terminator to about 800 or more for a fluorescent
dye-labeled ddNTP terminator may be associated
with potential three-dimensional structural

10

changes. These molecular alterations may
interfere with the process of incorporation of the
dye-labeled ddNTPs as chain terminators by the DNA
polymerase to the 3' end of an extending DNA
5   strand in terms of lowering the rate of
incorporation, lowering the processivity of the
enzyme for this new substrate, reducing the
enzyme-terminator binding specificity and changing
the enzyme-terminator binding kinetics.

10   For example, both Taq DNA polymerase and
Sequenase II™ (a T7 DNA polymerase) have been used
for radioisotope labeling DNA sequencing with
excellent results, and have been adapted for
fluorescent dye-labeled primer DNA sequencing.
15   But neither can be used for fluorescent dye-
labeled terminator DNA sequencing technologies.
As reported in U.S. Patent 5,614,365, when the Taq
DNA polymerase was used for fluorescent dye-
labeled terminator chemical reactions, the
20   reaction products generated no readable data on
the DNA sequencer. Most of the fluorescence was
either in unincorporated dye-ddNTPs at the leading
front of the test gel, or in fragments greater
than several hundred bases in length. Using a Taq
25   DNA polymerase mutant in which the amino acid,
phenylalanine, at position 667 of its amino acid
sequence has been replaced by a tyrosine and which
has an increased ability to incorporate
dideoxynucleotides (6,000 times more efficient),
30   to replace the unmodified Taq DNA polymerase for
the experiment, the results are significantly
improved. This F667Y mutant of Taq DNA polymerase
is now marketed by Amersham Life Science, Inc.
under the trademark ThermoSequenase™. It is used
35   for cycle-sequencing in which the enzymatic
reaction mixture is subjected to numerous cycles
of extension-termination, denaturing and annealing

11

to ensure that sufficient dye-terminator-labeled
enzymatic reaction products are generated for the
DNA sequencing procedure. Because of the low
processivity of the parent Taq DNA polymerase,

5   ThermoSequenase™ is not recommended for direct DNA
sequencing without precyclings. Like Taq DNA
polymerase, ThermoSequenase™ lacks a proof-reading
exonuclease activity.

     Bacillus stearothermophilus, *Bacillus*

10   *caldotenax* and *Bacillus caldolyticus* are
classified as mesophilic microbes; although their
DNA polymerases are referred to as thermostable
(most active at 65°C) they are inactivated at 70°C
or above. This is contrasted with other enzymes,

15   such as Taq, which are truly thermophilic--that
is, their DNA polymerases tolerate and remain
active at temperatures higher than 95°C. These
mesophilic bacillus strains, especially *Bacillus*
stearothermophilus, produce DNA polymerases that

20   are useful in DNA sequencing applications.
However, a disadvantage of the DNA polymerases of
these strains is that during DNA sequencing they
all exhibit a high degree of selective
discrimination against incorporation of certain

25   particular members of fluorescent dye-labeled
ddNTPs, namely the fluorescent dye-labeled ddCTP
and fluorescent dye-labeled ddATP, as terminators
onto the 3' end of the extending DNA fragments
during enzymatic reaction. This peculiar

30   characteristic of selective discrimination against
incorporation of fluorescent dye-labeled ddCTP and
ddATP of the natural DNA polymerases isolated from
Bacillus stearothermophilus and *Bacillus*
caldotenax was not previously recognized. Such

35   selective discrimination is apparently sequence-
related, and cannot be corrected or compensated by
mere adjustment of the concentrations of the

12

dNTPs.

Thus, there is a need for a mesophilic bacillus DNA polymerase that does not selectively discriminate against incorporation of fluorescent
5   dye-labeled ddCTP and ddATP, during dye terminator DNA sequencing.

SUMMARY OF THE INVENTION

This invention addresses the above-described problems associated with mesophilic
10   bacillus DNA polymerases by providing novel DNA polymerases which, during direct DNA sequencing, reduce the innate selective discrimination against the incorporation of fluorescent dye-labeled ddCTP and fluorescent dye-labeled ddATP, without
15   increasing the rate of incorporation of the other two dye-labeled ddNTP terminators (ddTTP and ddGTP) excessively. In particular, this invention provides a novel genetic modification of the amino acid sequence of a highly processive DNA
20   polymerase (such as isolated from *Bacillus stearothermophilus*, *Bacillus caldotenax* or *Bacillus caldolyticus*) that, unmodified, selectively discriminates against incorporation of fluorescent dye-labeled dideoxynucleotide
25   terminators ddATP and ddCTP (but does not discriminate against incorporation of fluorescent dye-labeled dideoxynucleotide terminators ddTTP and ddGTP). The modification results in a reduction of the innate selective discrimination
30   against incorporation of fluorescent dye-labeled dideoxynucleotide terminators ddATP and ddCTP, such that all four of the ddNTP terminators are effectively incorporated into the DNA primer elongated by the DNA polymerase. Thus, the
35   modified DNA polymerase of this invention is effective in reducing the innate selective

13

discrimination against incorporation of
fluorescent dye-labeled dideoxynucleotide
terminators ddATP and ddCTP characteristic of the
DNA polymerase in its unmodified state.

5   In particular, the preferred DNA
polymerase is a modification of a DNA polymerase
isolated from a strain of a mesophilic bacterium,
such as *Bacillus stearothermophilus*, *Bacillus
caldotenax* or *Bacillus caldolyticus*. The approach
10   of modifying the DNA polymerase described herein
may be used to modify other DNA polymerases which
share a close amino acid homology of a DNA
polymerase isolated from a strain *Bacillus
stearothermophilus*, *Bacillus caldotenax* or
15   *Bacillus caldolyticus*, as long as the unmodified
DNA polymerases have a selective discrimination
against incorporation of fluorescent dye-labeled
dideoxynucleotide ddCTP and/or ddATP as
terminators in the enzymatic reaction for
20   preparing materials for automated fluroescent DNA
sequencing. Consequently, it is preferred that
the modified DNA polymerase has an amino acid
sequence that shares not less than 95% homology of
a DNA polymerase isolated from a strain of
25   *Bacillus stearothermophilus*, *Bacillus caldotenax*
or *Bacillus caldolyticus*.

          The particularly preferred mesophilic
species is *Bacillus stearothermophilus*, which is
highly heterogeneous. This is indicated by the
30   wide range of DNA base compositions as well as the
range of the phenotypic properties of strains
assigned to this species (see Bergey's Manual of
Systemic Bacteriology, Eds. P.H.A. Sneath, N.S.
Mair, M.E. Sharpe and J.G. Holt, Williams &
35   Wilkins, 1986, Vol. 2, page 1135). Therefore, it
is reasonable to assume that the amino acid
sequences of DNA polymerases isolated from various

strains would be heterogeneous with potential
functional differences. Although DNA polymerases
isolated from the known standard strains of
Bacillus stearothermophilus have been shown to

5   lack a 3'-5' exonuclease activity, a questionable
trace of "contaminating" 3'5' exonuclease has been
observed in a purified DNA polymerase preparations
(see Kaboev et al., J. Bacteriology, Vol. 145,
page 21-26, 1981).

10   Consequently, the inventors began to
address the above-identified problems in the art
by discovering a strain of *Bacillus
stearothermophilus* (designated strain No. 320 for
identification purposes; described in U.S. Patent

15   5,747,298) that produces a DNA polymerase
(designated Bst 320) with a proof-reading 3'-5'
exonuclease activity which is absent in DNA
polymerases isolated from other strains of
Bacillus stearothermophilus. (For this invention,

20   the term "proof-reading" is intended to denote
that the DNA polymerase is capable of removing
mismatched nucleotides from the 3' terminus of a
newly formed DNA strand at a faster rate than the
rate at which nucleotides correctly matched with

25   the nucleotides of the template are removed during
DNA sequencing.) The strain Bst 320 was deposited
on October 30, 1995 in the American Type Culture
Collection, located at 12301 Parklawn Drive,
Rockville, Maryland 20852, and has been given ATCC

30   Designation No. 55719. The DNA polymerase
isolated from Bst 320 is composed of 587 amino
acids as are the DNA polymerases of other known
strains of *Bacillus stearothermophilus*, such as,
for instance, the strains deposited by Riggs et al

35   (Genbank Accession No. L42111) and by Phang et al.
(Genbank Accession No. U23149). However, the Bst
320 shares only 89.1% sequence identity at protein

level with the *Bacillus stearothermophilus* DNA
polymerase deposited by Riggs et al., and shares
only 87.4% sequence identity at protein level with
the *Bacillus stearothermophilus* DNA polymerase

5   deposited by Phang et al. For comparison, the
above-referenced enzyme deposited by Riggs et al.
and the enzyme deposited by Phang et al. share
96.9% of their amino acid sequence identity.

The inventors studied a thermostable DNA

10   polymerase isolated from a different species,
Bacillus caldotenax (Bca), which also has an
optimum active temperature at 65°C. The inventors
discovered that the Bst 320 DNA polymerase shares
88.4% of the amino acid sequence identity with Bca

15   DNA polymerase (Uemori et al. J. Biochem. 113:
401-410, 1993). Based on homology of the amino
acid sequences, Bst 320 DNA polymerase is as close
to DNA polymerases isolated from *Bacillus
stearothermophilus* as to the DNA polymerase

20   isolated from *Bacillus caldotenax*, i.e. another
species of bacillus. It was also discovered that
both Bst 320 DNA polymerase and Bca DNA polymerase
functionally exhibit 3'-5' exonuclease activity,
which is not associated with known amino acid

25   sequence exonuclease motifs I, II and III as in
the *E. coli* DNA polymerase I model, or other known
Bacillus stearothermophilus polymerases.

The inventors has studied the DNA
polymerases of three different strains of *Bacillus

30   stearothermophilus* (including DNA polymerase
obtained from Bst 320) and the DNA polymerase of
Bacillus caldotenax and found that they all
exhibit a high degree of selective discrimination
against incorporation of certain particular

35   members of fluorescent dye-labeled ddNTPs, namely
the fluorescent dye-labeled ddCTP and fluorescent
dye-labeled ddATP, as terminators onto the 3' end

16

of the extending DNA fragments during enzymatic reaction. This is especially the case when the preceding 3' end base of the extending DNA fragment is a dGMP (G) or a dAMP (A). (By "dNTP" it is intended to denote the four commonly known deoxynucleotide triphosphates, dATP, dTTP, dCTP, and dGTP.)

This selective discrimination causes missing peaks and ambiguous peaks on a color plot generated by the automated fluorescent DNA sequencer, and causes loss of sequencing data and erroneous base callings. This is shown in Figures 6 and 8.

This disadvantage of the natural bacillus DNA polymerases in fluorescent dye-labeled terminator DNA sequencing cannot be corrected or compensated by mere adjustment of the concentrations of the dNTPs and the fluorescent dye-labeled ddNTPs in the reaction mixture. This selective discrimination against the specific dye-labeled ddNTPs is also sequence-related as demonstrated with respect to Bst in Figures 6 and 8, in which the missing or ambiguous "C" peaks and "A" peaks tend to occur immediately following a preceding "G" peak or a preceding "A" peak. Of particular interest is the fact that the "C" and "A" peaks immediately following a preceding "C" or a preceding "T" peak are quite strong and resolvable in the same color plot analysis, indicating that the concentrations of dNTPs and the fluorescent dye-labeled ddCTP and the fluorescent dye-labeled ddATP were adequate for the termination reaction.

According to the structural model studies carried out on *E. coli* DNA polymerase I (Klenow fragment), certain amino acids in a particular region or regions of a DNA polymerase

appear to play important roles in dNTP and ddNTP
bindings and their final incorporation, and affect
discrimination between deoxy and dideoxynucleotide
substrates. For example, mutation of the amino

5   acids arginine, asparagine, lysine, tyrosine,
phenylalanine, aspartate, and glutamate in certain
locations of amino acid sequences of Klenow
fragment may affect the binding of dNTP and
discrimination between deoxy and

10   dideoxynucleotides. (See: Joyce, C.M., Current
Opinion in Structural Biology, 1:123-129, 1991.
Joyce and Steitz, Annu. Rev. Biochem., 63:777-822,
1993, page 800. Carrol et al., Biochemistry
30:804-813, 1991).

15   The problem which faced the inventors
was how to reduce the selective discrimination
against the incorporation of fluorescent dye-
labeled ddCTP and fluorescent dye-labeled ddATP by
site-directed mutagenesis of a DNA polymerase,

20   without increasing the rate of incorporation of
the other two dye-labeled ddNTP terminators
excessively. In particular, the new mutant must
be able to incorporate more correctly base-matched
dye-labeled ddCTP and/or dye-labeled ddATP

25   terminators to the dGMP (G) and dAMP (A) bases,
than to the dCMP (C) and dTMP (T) bases of the
extending DNA fragments during enzymatic reaction.
A blanket increase in the ability of an enzyme to
incorporate all four dye-labeled ddNTPs to the

30   same proportion would serve no useful purpose for
the group of DNA polymerases isolated from
mesophilic bacilli since, unlike the Taq DNA
polymerase, the unmodified natural enzymes of
Bacillus stearothermophilus and *Bacillus*

35   *caldotenax* already possess a high ability to
incorporate fluorescent dye-labeled ddGTP and
fluorescent dye-labeled ddTTP, and even the

18

fluorescent dye-labeled ddCTP and dye-labeled
ddATP provided at the immediately preceding base
at the 3'end of the extending DNA fragment is not
a "G" or an "A".

5   The inventors found that DNA polymerases
isolated from strains of *Bacillus*
stearothermophilus and *Bacillus caldotenax* possess
the same amino acids at certain specific positions
in their amino acid sequence. For example, they

10   all have leucine-glutamate-glutamate at positions
corresponding to positions 342-344 and
phenylalanine at a position
corresponding to position 422 of the amino acid
sequence of the DNA polymerase isolated from No

15   320 strain of *Bacillus stearothermophilus*. The
inventors further discovered that the most optimal
modification to solve the problem of selective
discrimination in direct fluorescent DNA
sequencing for these DNA polymerases is to modify

20   the four amino acids of the natural DNA
polymerases referenced above in such a form that
threonine-proline-leucine substitute respectively
for leucine-glutamate-glutamate at positions 342-
344 and tyrosine substitutes for phenylalanine at

25   position 422 in their amino acid sequences.
Accordingly, the nucleotide sequence encoding the
natural forms of the DNA polymerases are modified
at positions 1024-1032 from CTCGAAGAG to ACCCCACTG
and at position 1265 from T to A to encode for the

30   DNA polymerases having the desired properties.
The combined effects of these amino acid
modifications reduce the selective discrimination
against incorporation of fluorescent dye-labeled
ddCTP and dye-labeled ddATP of the naturally-

35   occurring mesophilic bacillus DNA polymerases
during enzymatic reaction for direct automated
fluorescent DNA sequencing.

19

Initially, the DNA polymerases used in
the inventors' research were obtained by
overexpression of the genes encoding the
naturally-occurring enzymes of *Bacillus*
5   *stearothermophilus* and *Bacillus caldotenax*.
Subsequently, modified DNA polymerases obtained by
overexpression of the site-directed mutated genes
were used. This invention provides both the
nucleotide and amino acid sequence for a modified
10   DNA polymerase to illustrate the practice of this
new approach of modifying a special group of DNA
polymerases, as described below.

In one preferred embodiment, the Bst 320
DNA polymerase is used for the unmodified,
15   naturally-occurring DNA polymerase, although DNA
polymerases isolated from other strains of
mesophilic bacilli (for instance, *Bacillus*
stearothermophilus and *Bacillus caldotenax*) can be
used as the starting enzymes for the genetic
20   modification. As noted above, the Bst 320 DNA
polymerase is also capable of proofreading 3'-5'
exonuclease activity. In particular, the
invention provides the DNA and amino acid
sequences for the isolated and purified DNA
25   polymerase having this function. These sequences
are also described below.

The invention also contemplates an
isolated strain of *Bacillus stearothermophilus*
which produces a DNA polymerase having an ability
30   to reduce selective discrimination against
incorporation of fluorescent dye-labeled
dideoxynucleotide terminators ddCTP and ddATP, but
not fluorescent dye-labeled dideoxynucleotide
terminators ddGTP and ddTTP, in the presence of
35   dNTPs and the four fluorescent dye-labeled
dideoxynucleotide terminators. Preferably, the Bst
strain produces a DNA polymerase which also has

20

proofreading 3'-5' exonuclease activity during DNA
sequencing of a DNA strand from a template.

As mentioned above, the invention also
contemplates DNA polymerases obtained or otherwise
5   derived from any bacillus strain, or made
synthetically, as long as the amino acid sequences
of the naturally-occurring DNA polymerases have
leucine-glutamate-glutamate at positions
corresponding respectively to positions 342-344 of
10   Bst 320 DNA polymerase and phenylalanine at a
position corresponding to position 422 of Bst 320
DNA polymerase. For example, DNA polymerases
derived from other strains of *Bacillus*
stearothermophilus or *Bacillus caldotenax* or other
15   mesophilic bacilli may be easily modified using
conventional DNA modification techniques to
include the amino acid or nucleotide substitutions
identified above.

The invention also provides a DNA
20   construct comprising at least one of the above-
described DNA polymerase sequences and a vector
(such as a cloning vector or an expression
vector), for introducing the DNA construct into
eucaryotic or procaryotic host cells (such as an
25   *E. coli* host cell). In addition, the invention
further provides a host cell stably transformed
with the DNA construct in a manner allowing
production of the peptide encoded by the DNA
segment in the construct.

30   The invention also provides improved
methods for replicating DNA and sequencing DNA
using the above-described DNA polymerases of the
invention. The DNA polymerases are useful in both
direct dye terminator DNA sequencing and dye-
35   primer DNA sequencing.

Preferably, the method of sequencing a
DNA strand may comprise the steps of:

21

    i)   hybridizing a primer to a DNA template to be sequenced;

    ii)   extending the primer using a DNA polymerase which has an ability to reduce selective discrimination against incorporation of fluorescent dye-labeled dideoxynucleotide terminators ddCTP and ddATP, in the presence of adequate amounts of nucleotide bases dATP, dGTP, dCTP and dTTP, or their analogs, and the four fluorescent dye-labeled dideoxynucleotide terminators,

    under such conditions that the DNA strand is sequenced.

    Further objects and advantages of the invention will become apparent from the description and examples below.

BRIEF DESCRIPTION OF THE DRAWINGS

    In the Figures and throughout this disclosure, "HiFi Bst" or "Bst 320" DNA polymerase refers to the unmodified naturally occurring DNA polymerase having proofreading 3'-5' exonuclease activity, either isolated from the cells of No. 320 strain of *Bacillus stearothermophilus* or produced by overexpression of the gene encoding this naturally occurring DNA polymerase. (This Bst strain and DNA polymerase are described in U.S. Patent 5,747,298.) "HiFi Bst-II" refers to the modified form of "HiFi Bst" DNA polymerase which has an ability to reduce selective discrimination against fluorescent dye-labeled ddCTP and ddATP. HiFi Bst-II is an example of one preferred embodiment of this invention.

    Figure 1. This graph shows the thermostability at 65°C of HiFI Bst-II and HiFI Bst.

22

Y: relative polymerase activity (%)

X: incubation time (minutes).

Figure 2. This shows a autoradiograph of a DNA sequencing gel obtained by using radiolabeled primer with HiFi Bst-II and HiFi Bst, and shows the dideoxy-nucleotide incorporation of HiFi Bst-II and HiFi Bst in a reaction mixture with a suboptimally low ddNTP/dNTP ratios.

Template: single-stranded M13mp18;

Primer: -20M13 forward primer.

Figure 3. This shows a autoradiograph of a DNA sequencing gel obtained by using radiolabeled dATP with HiFi Bst and HiFi Bst-II in reaction mixtures with optimized ddNTP/dNTP ratios. The sequence pattern with HiFi Bst-II is better than that with HiFi Bst.

Template: single-stranded M13mp18;

Primer: -20M13 forward primer.

Figure 4. This shows the results of dye-primer DNA sequencing with HiFi Bst

Template: single-stranded pGEM-3Zf(+);

Primer: -21M13 forward DYEnamic Energy Transfer Dye Primers.

Figure 5. This shows the results of dye-primer DNA sequencing with HiFi Bst-II.

Template: single-stranded M13mp18;

Primer: -21M13 forward DYEnamic Energy Transfer Dye Primers.

Figure 6. This shows the results of dye-terminator DNA sequencing with HiFi Bst

Template: single-stranded pGEM-3Zf(+);

Primer: -20M13 forward primer.

Figure 7. This shows the results of dye-terminator DNA sequencing with HiFi Bst-II.

Template: single-stranded M13mp18;

Primer: -20M13 forward primer.

Figure 8. Like Figure 6, this shows the
results of four fluorescent dye-labeled
terminators DNA sequencing with HiFi Bst. In
Figure 8 corrections of the missing or ambiguous

5   bases, according to the known pGEM sequence, are
indicated below the letters "N" or below the
incorrect base letters.

Template: single-stranded pGEM-
3Zf(+);

10   Primer: -20M13 forward primer.

DETAILED DESCRIPTION OF THE INVENTION

The DNA polymerases of the invention are
capable of reducing selective discrimination
against incorporation of fluorescent dye-labeled

15   dideoxynucleotide terminators ddCTP and ddATP,
(but not ddGTP and ddTTP), in the presence of
adequate amounts of dNTPs and the four
terminators.

The inventors discovered that certain

20   modifications of the amino acid sequence of DNA
polymerases (i.e., modifying the amino acids at
positions 342-344 to substitute threonine, proline
and leucine, respectively, for leucine, glutamate
and glutamate, and modifying the amino acid at

25   position 422 to substitute tyrosine for
phenylalanine, as corresponding to the amino acid
sequence of Bst 320 DNA polymerase) result in a
marked reduction of the innate selective
discrimination against incorporation of

30   fluorescent dye-labeled dideoxynucleotide ddCTP
and ddATP, that is characteristic of the
unmodified DNA polymerase. This reduction of
selective discrimination can be demonstrated by
direct automated fluorescent terminator DNA

35   sequencing as recovered missing or ambiguous "C"

24

peaks and/or "A" peaks in the automated sequencing
results, using a known template as standard.

The preferred source for the DNA
polymerase is *Bacillus stearothermophilus* because
5   DNA polymerase isolated from this mesophilic
bacillus is highly processive, can be used and
stored in dried down form, has an optimum
temperature at 65°C, and can be used for direct
DNA sequencing without precycling. The preferred
10   Bst DNA polymerase is that isolated from strain
320 with an amino acid sequence as follows:

Amino acid sequence (SEQ ID NO:2):
AEGEKPLEEM EFAIVDVITE EMLADKAALV VEVMEENYHD
APIVGIALVN EHGRFFMRPE TALADSQFLA WLADETKKKS
15   MFDAKRAVVA LKWKGIELRG VAFDLLLAAY LLNPAQDAGD
IAAVAKMKQY EAVRSDEAVY GKGVKRSLPD EQTLAEHLVR
KAAAIWALEQ PFMDDLRNNE QDQLLTKLEH ALAAILAEME
FTGVNVDTKR LEQMGSELAE QLRAIEQRIY ELAGQEFNIN
SPKQLGVILF EKLQLPVLKK TKTGYSTSAD VLEKLAPHHE
20   IVENILHYRQ LGKLQSTYIE GLLKVVRPDT GKVHTMFNQA
LTQTGRLSSA EPNLQNIPIR LEEGRKIRQA FVPSEPDWLI
FAADYSQIEL RVLAHIADDD NLIEAFQRDL DIHTKTAMDI
FQLSEEEVTA NMRRQAKAVN FGIVYGISDY GLAQNLNITR
KEAAEFIERY FASFPGVKQY MENIVQEAKQ KGYVTTLLHR
25   RRYLPDITSR NFNVRSFAER TAMNTPIQGS AADIIKKAMI
DLAARLKEEQ LQARLLLQVH DELILEAPKE EIERLCELVP
EVMEQAVTLR VPLKVDYHYG PTWYDAK

The characters represent the following amino
acids:
30   where,
A: alanine (Ala)   M: methionine (Met)
C: cysteine (Cys)   N: asparagine (Asn)
D: aspartic acid (Asp)   P: proline (Pro)
E: glutamic acid (Glu)   Q: glutamine (Gln)

25

F: phenylanaline (Phe)   R: arginine (Arg)
G: glycine (Gly)   S: serine (Ser)
H: histidine (His)   T: threonine (Thr)
I: isoleucine (Ile)   V: valine (Val)
5   K: lysine (Lys)   W: tryptophan (Trp)
L: leucine (Leu)   Y: tyrosine (Tyr)

The Bst 320 DNA polymerase is characterized by possessing a proofreading 3'-5' exonuclease activity.

10   The nucleotide sequence encoding the unmodified Bst 320 DNA polymerase is indicated in SEQ ID NO:1, in Example 2 below.

The following amino acid sequence represents the modified Bst 320 DNA polymerase as
15   the preferred embodiment of this invention, modified from the naturally-occurring Bst 320 DNA polymerase at positions 342-344 to substitute threonine, proline and leucine, respectively, for leucine, glutamate and glutamate, and at position
20   422 to substitute tyrosine for phenylalanine.

Amino acid sequence (SEQ ID:No 4):
MAEGEKPLEEMEFAIVDVITEEMLADKAALVVEVMEENYHDAPIV
GIALVNEHGRFFMRPETALADSQFLAWLADETKKKSMFDAKRAVV
ALKWKGIELRGVAFDLLLAAYLLNPAQDAGDIAAVAKMKQYEAVR
25   SDEAVYGKGVKRSLPDEQTLAEHLVRKAAAIWALEQPFMDDLRNN
EQDQLLTKLEHALAAILAEMEFTGVNVDTKRLEQMGSELAEQLRA
IEQRIYELAGQEFNINSPKQLGVILFEKLQLPVLKKTKTGYSTSA
DVLEKLAPHHEIVENILHYRQLGKLQSTYIEGLLKVVRPDTGKVH
TMFNQALTQTGRLSSAEPNLQNIPIRTPLGRKIRQAFVPSEPDWL
30   IFAADYSQIELRVLAHIADDDNLIEAFQRDLDIHTKTAMDIFQLS
EEEVTANMRRQAKAVNYGIVYGISDYGLAQNLNITRKEAAEFIER
YFASFPGVKQYMENIVQEAKQKGYVTTLLHRRRYLPDITSRNFNV
RSFAERTAMNTPIQGSAADIIKKAMIDLAARLKEEQLQARLLLQV
HDELILEAPKEEIERLCELVPEVMEQAVTLRVPLKVDYHYGPTWY
35   DAK

26

The underlined amino acids are substituted amino
acids produced by site-directed mutation of the
naturally-occurring Bst 320 DNA polymerase.

5   The modified Bst 320 DNA polymerase is
encoded by a DNA sequence such as the following
(SEQ ID NO:3):

```
                                                 ATG
      GCCGAAGGGG AGAAACCGCT TGAGGAGATG GAGTTTGCCA
      TCGTTGACGT CATTACCGAA GAGATGCTTG CCGACAAGGC
10    AGCGCTTGTC GTTGAGGTGA TGGAAGAAAA CTACCACGAT
      GCCCCGATTG TCGGAATCGC ACTAGTGAAC GAGCATGGGC
      GATTTTTTAT GCGCCCGGAG ACCGCGCTGG CTGATTCGCA
      ATTTTTAGCA TGGCTTGCCG ATGAAACGAA GAAAAAAAGC
      ATGTTTGACG CCAAGCGGGC AGTCGTTGCC TTAAAGTGGA
15    AAGGAATTGA GCTTCGCGGC GTCGCCTTTG ATTTATTGCT
      CGCTGCCTAT TTGCTCAATC CGGCTCAAGA TGCCGGCGAT
      ATCGCTGCGG TGGCGAAAAT GAAACAATAT GAAGCGGTGC
      GGTCGGATGA AGCGGTCTAT GGCAAAGGCG TCAAGCGGTC
      GCTGCCGGAC GAACAGACGC TTGCTGAGCA TCTCGTTCGC
20    AAAGCGGCAG CCATTTGGGC GCTTGAGCAG CCGTTTATGG
      ACGATTTGCG GAACAACGAA CAAGATCAAT TATTAACGAA
      GCTTGAGCAC GCGCTGGCGG CGATTTTGGC TGAAATGGAA
      TTCACTGGGG TGAACGTGGA TACAAAGCGG CTTGAACAGA
      TGGGTTCGGA GCTCGCCGAA CAACTGCGTG CCATCGAGCA
25    GCGCATTTAC GAGCTAGCCG GCCAAGAGTT CAACATTAAC
      TCACCAAAAC AGCTCGGAGT CATTTTATTT GAAAAGCTGC
      AGCTACCGGT GCTGAAGAAG ACGAAAACAG CTATTCGAC
      TTCGGCTGAT GTGCTTGAGA AGCTTGCGCC GCATCATGAA
      ATCGTCGAAA ACATTTTGCA TTACCGCCAG CTTGGCAAAC
30    TGCAATCAAC GTATATTGAA GGATTGTTGA AGTTGTGCG
      CCCTGATACC GGCAAAGTGC ATACGATGTT CAACCAAGCG
      CTGACGCAAA CTGGGCGGCT CAGCTCGGCC GAGCCGAACT
      TGCAAAACAT TCCGATTCGG ACCCCACTGG GGCGGAAAAT
      CCGCCAAGCG TTCGTCCCGT CAGAGCCGGA CTGGCTCATT
35    TTCGCCGCCG ATTACTCACA AATTGAATTG CGCGTCCTCG
      CCCATATCGC CGATGACGAC AATCTAATTG AAGCGTTCCA
      ACGCGATTTG GATATTCACA CAAAAACGGC GATGGACATT
```

```
TTCCAGTTGA GCGAAGAGGA AGTCACGGCC AACATGCGCC
GCCAGGCAAA GGCCGTTAAC TACGGTATCG TTTACGGAAT
TAGCGATTAC GGATTGGCGC AAAACTTGAA CATTACGCGC
AAAGAAGCTG CCGAATTTAT CGAACGTTAC TTCGCCAGCT
TTCCGGGCGT AAAGCAGTAT ATGGAAAACA TAGTGCAAGA
AGCGAAACAG AAAGGATATG TGACAACGCT GTTGCATCGG
CGCCGCTATT TGCCTGATAT TACAAGCCGC AATTTCAACG
TCCGCAGTTT TGCAGAGCGG ACGGCCATGA ACACGCCAAT
TCAAGGAAGC GCCGCTGACA TTATTAAAAA AGCGATGATT
GATTTAGCGG CACGGCTGAA AGAAGAGCAG CTTCAGGCTC
GTCTTTTGCT GCAAGTGCAT GACGAGCTCA TTTTGGAAGC
GCCAAAAGAG GAAATTGAGC GATTATGTGA GCTTGTTCCG
GAAGTGATGG AGCAGGCCGT TACGCTCCGC GTGCCGCTGA
AAGTCGACTA CCATTACGGC CCAACATGGT ATGATGCCAA
A
```

The characters represent the following
nucleotides:

    A: Adenosine   T: Thymidine
    C: Cytidine   G: Guanosine

        The underlined nucleotides are
substituted nucleotides produced by site-directed
mutation of the naturally-occurring Bst 320
polymerase. (As would be apparent to someone
skilled in this art, this DNA sequence does not
indicate the starting codon.)

        The invention also contemplates any DNA
sequence that is complementary to the modified Bst
320 DNA sequence, for instance, DNA sequences that
would hybridize to the above DNA sequence of the
modified DNA polymerase under stringent
conditions. As would be understood by someone
skilled in the art, the invention also
contemplates any DNA sequence that encodes a
peptide having these characteristics and
properties (including degenerate DNA code).

        This invention also contemplates allelic
variations and mutations (for instance, adding or

28

deleting nucleotide or amino acids, sequence
recombination or replacement or alteration) which
result in no substantive change in the function of
the DNA polymerase or its characteristics. For

5   instance, the DNA polymerases encompass non-
critical substitutions of nucleotides or amino
acids that would not change functionality (i.e.,
such as those changes caused by a transformant
host cell). In addition, the invention is

10   intended to include fusion proteins and muteins of
the unique DNA polymerases of this invention.

The DNA sequences and amino acid
sequences for the modified DNA polymerase of this
invention are also obtainable by, for instance,

15   isolating and purifying DNA polymerase from a
Bacillus stearothermophilus, or a bacterial strain
otherwise derived from *Bacillus
stearothermophilus*, or other mesophilic bacillus
strains such as *Bacillus caldotenax* or *Bacillus*

20   *caldolyticus*. The DNA polymerases obtained from
these organisms may be easily modified using
conventional DNA modification techniques to
achieve the reduction in fluorescent dye-labeled
ddCTP and ddATP selective discrimination, as long

25   as the unmodified amino acid sequences have
leucine-glutamate-glutamate at positions
corresponding respectively to positions 342-344 of
Bst 320 DNA polymerase and phenylalanine at a
position corresponding to position 422 of Bst 320

30   DNA polymerase. For instance, using the primers
and methods of screening described herein, someone
skilled in the art could isolate a DNA polymerase
having the same properties and function from other
strains.

35   In the DNA polymerases currently used in
conventional DNA sequencing protocols, it is
preferred that the enzymes have low or no

29

exonuclease activity. However, in this invention,
it is preferred that the DNA polymerases have a
function of high fidelity ("HiFi") nucleotide
incorporation. Therefore, in one preferred

5   embodiment the invention entails modification of a
naturally-occurring Bst DNA polymerase having a
proofreading 3'-5' exonuclease activity. This
preferred modified DNA polymerase (e.g., "HiFi
Bst-II") has a nucleotide sequence indicated in

10   SEQ ID:NO 3 and an amino sequence indicated in SEQ
ID:NO 4. To initially obtain a Bst DNA polymerase
having proofreading 3'-5' activity, strains of
Bacillus stearothermophilus can be segregated into
different groups according to the proof-reading

15   exonuclease activity of their respective DNA
polymerases.

The invention also provides a DNA
construct comprising at least one of the DNA
sequences of the modified DNA polymerase and a

20   vector (such as a cloning vector or an expression
vector), for introducing the DNA construct into
host cells. An example of a suitable vector is
pYZ23/LF, described below.

The host cells need only be capable of

25   being stably transformed with the DNA construct in
a manner allowing production of the peptide
encoded by the DNA segment in the construct
(preferably in large quantity). The host cells
may be of eucaryotic or procaryotic origin (such

30   as a E. coli host cell). For instance, the host
cell may be a mesophilic organism, although this
is not a necessary requirement in order that a
host cell be effective.

The invention also provides improved

35   methods for DNA sequencing using the above-
described novel DNA polymerases. The methods
entail sequencing a DNA strand by conventional

protocols with the following modifications:

     i)   hybridizing a primer to a DNA template to be sequenced;

     ii)   extending the primer using a DNA polymerase described above, in the presence of radiolabeled dATP, nucleotides dGTP, dCTP and dTTP, or their analogs, and ddNTP chain terminators; and

     iii)   allowing a DNA strand to be sequenced.

     All four dNTPs, including dCTP, are incorporated equally effectively in the chain elongation during sequencing reaction catalyzed by the DNA polymerases of the invention with a high processivity and a high elongating rate.

     Preferably the nucleotide premix concentrations of modified Bst DNA polymerase used in radiolabeled DNA sequencing are as following:

A mix: dATP 0.8μM, dCTP 80μM, dGTP 80μM, dTTP 80μM, ddATP 25μM;

C mix: dATP 0.8μM, dCTP 8μM, dGTP 80μM, dTTP 80μM, ddCTP 20μM;

G mix: dATP 0.8μM, dCTP 80μM, dGTP 8μM, dTTP 80μM, ddGTP 50μM;

T mix: dATP 0.8μM, dCTP 80μM, dGTP 80μM, dTTP 8μM, ddTTP 50μM. (This mixture is useful for the particular modified Bst 320 DNA polymerase set forth above, as well as for other modified Bst DNA polymerases.)

     In addition, the invention contemplates other uses of the modified DNA polymerases. For instance, the DNA polymerase can also be use in (1) filling-in 5' overhangs of DNA fragments; (2) synthesis of DNA probes by random primers labeling methodology; and (3) site-directed mutagenesis.

     The following non-limiting examples are

31

illustrative of the invention.

Example 1: Screening for Bst polymerases

        This invention also involves a method to
measure the proof-reading 3'-5' exonuclease
5   activity of purified DNA polymerases. The method
is useful to screen a large number of bacterial
strains, such as *Bacillus stearothermophilus* and
other mesophilic bacterial strains, to select a
strain which produces a DNA polymerase with a high
10   proof-reading 3'-5' exonuclease activity. For
instance, the method to test the proof-reading 3'-
5' exonuclease activity of DNA polymerase was
carried out as follows.

        A DNA primer and two DNA templates with
15   following sequences were synthesized chemically,
using a DNA synthesizer.

17-base primer   5' CATTTTGCTGCCGGTCA 3'
                            1 mg/ml
                        (SEQ ID NO:5)

20   Template (a)   3'------GTAAAACGACGGCCAGTCTT------5'
                              10 mg/ml
                           (SEQ ID NO:6)

        Template (b)   3'-----GTAAAACGACGGCCAGTCGG-----5'
                               10 mg/ml
25   (SEQ ID NO:7)

        To produce the radiolabeled primer, 1 μl
(1 μg) of primer, 5 μl (50 μg) of template (a), 1
μl of [α-^{32}P] dATP (800 Ci/mmole), 1 μl of dGTP
(0.5 mM), 1 μl of Taq DNA polymerase (1 unit), and
30   1 μl of buffer consisting of 500 mM Tris-Cl, pH
9.0, and 150 mM MgCl$_2$, were mixed in a test tube
and incubated in a 65°C water bath for 5 minutes.
The mixture was subject to alkaline denaturing gel

electrophoresis. The radioactive band containing
the 20-base nucleotide was isolated and dissolved
in 12 µl of 10 mM Tris-Cl buffer, containing 1 mM
EDTA, pH 8.0. The final product represents the

5   following labeled 20-base primer.

       5'   CATTTTGCTGCCGGTCAGA*A*   3'

                          (* = ^{32}P labeled)

                  (SEQ ID NO:8)

          To produce radiolabeled primer-template

10   complexes, 5 µl of the labeled primer was mixed
with 10 µl of template (a) or template (b)
respectively to form the following:

Complex (a)
5'CATTTTGCTGCCGGTCAGA*A* 3' (same as SEQ ID NO:8)

15   3'GTAAAACGACGGCCAGTCT T 5' (same as SEQ ID NO:6)

Complex (b)
5' CATTTTGCTGCCGGTCAGA*A* 3' (same as SEQ ID NO:8)
3' GTAAAACGACGGCCAGTCG G 5' (same as SEQ ID NO:7)

The free radiolabeled primer was removed through a

20   G-50 Sephadex column.

          An aliquot of complex (a) which had two
correctly matched radiolabeled A*s at the 3'
terminus of the primer, and an aliquot of complex
(b) which had two mismatched A*s at the 3'

25   terminus of the primer, were then pipetted into
two individual vials of scintillation fluid and
their radioactivity was measured in a
scintillation counter, and both complexes were
adjusted with buffer to a concentration containing

30   the same molarity of incorporated [α-^{32}P] dAMP.

          To perform the proof-reading 3'-5'
exonuclease activity, 20 µl of complex (a) or
complex (b), 8 µl reaction buffer consisting of 15

mM Tris-Cl and 15 mM MgCl$_2$, pH 8.5, 4 units of DNA
polymerase, and enough water to make up a total
volume of 40 μl were pipetted into a test tube and
mixed well. The mixture was subdivided into

5   aliquots of 3 μl each in 0.5 ml microcentrifuge
tubes and was then covered with 3 μl paraffin in
each tube. The microcentrifuge tubes were
incubated in a 65°C water bath. At 1, 2, 3, 5, 10,
and 20 minutes, a pair of the microcentrifuge

10   tubes were taken out from the water bath and the
content of each tube was dotted onto a DE-81
Whatman filter paper. One of each pair of the
filter papers was put in scintillation fluid
directly and the radioactivity was counted in cpm

15   value in a scintillation counter; the other was
washed three times in 0.3 M sodium phosphate
buffer, pH 6.8 before being put into the
scintillation fluid for counting.

        The difference in radioactivity

20   expressed in cpm value between the washed filter
paper and the unwashed filter paper in each pair
was interpreted as representing the relative
quantity of labeled nucleotides excised by the 3'-
5' exonuclease activity from the 3' terminus of

25   the primer. A DNA polymerase that excised the
radiolabeled nucleotides A*s from complex (b) more
efficiently than from complex (a) possessed proof-
reading 3'-5'exonuclease activity. A DNA
polymerase that excised the radiolabeled

30   nucleotides A*s from complex (a) faster than from
complex (b), or at nearly the same rate, was
interpreted as possessing a non-specific 3'-5'
exonuclease activity which is considered
unsuitable for DNA sequencing.

35   Using these methods, a strain of
bacteria was isolated from among the strains of
Bacillus stearothermophilus from various sources

34

which is distinguished in its fast growth rate.
This strain reached an optimum exponential growth
within 3 hours for DNA polymerase production. The
strain was also able to produce a DNA polymerase
5   with a proof-reading 3'-5' exonuclease activity.
This strain of *Bacillus stearothermophilus* was
labeled Bst No. 320.

       As persons skilled in this art would
appreciate, the bacterial strain, or even the
10   strain of *Bacillus stearothermophilus*, from which
a mesophilic DNA polymerase of the invention can
be obtained may be derived using the above-
described methods or others known in the art from
strains of *Bacillus stearothermophilus* or other
15   bacteria strains (especially mesophilic strains),
including wild strains or mutant strains acquired
by various means, including spontaneous mutation.

       To prepare the preferred purified Bst
DNA polymerase, the cells of Bst No. 320 were
20   grown at 55°C in a liquid medium consisting of 1%
polypeptone, 0.5% yeast extract and 0.5% NaCl,
pH7.0-7.2. The 3 hr old cells were collected after
centrifugation and suspended in 4 volumes of TME
buffer (50mM Tris-HCl, pH7.5, 10mM β-
25   mercaptoethanol, and 2 mM EDTA), containing 100 mg
lysozyme and 23 mg phenylmethylsulphonyl-
fluoride/ml. The cells were broken by sonication
in ice. The supernatant was pooled after
centrifugation at 28,000 rpm in a Spinco L 30
30   rotor.

       The purified Bst DNA polymerase of the
invention was prepared according to Okazaki and
Kornberg (7) with appropriate slight modifications
and the large fragment of the DNA polymerase was
35   obtained by partial digestion of the whole DNA
polymerase with the proteinase subtilisin (type
Carlsberg) basically according to Jacobsen et al.

(8).

The procedure for purification of enzyme was followed as described in Ye and Hong (4). This Bst DNA polymerase possessed a proof-reading

5   3'-5' exonuclease activity.

The Bst polymerase was tested for proof-reading and non-specific 3'-5'exonuclease activities as described above. The results showed that the polymerase excised the mismatched

10   incorporated nucleotides from the 3' terminus of a double-stranded DNA at a high rate, reaching the plateau of hydrolysis in about 3 minutes, about 8 times more efficiently in the first 3 minutes of reaction than those correctly matched with the

15   nucleotides of the template. This enzyme is referred to herein as HiFi Bst DNA polymerase, and is distinguishable from Bst DNA polymerases isolated from other strains of *Bacillus stearothermophilus*.

20   This prodedure of using HiFi Bst as the DNA polymerase in the classic radiolabeling Sanger reaction for DNA sequencing and its autoradiograph are illustrated in Example 8. The data obtained by adapting this procedure to use HiFi Bst as the

25   DNA polymerase in dye-labeled primer automated fluorescent DNA sequencing are illustrated in Example 9. These results indicate that HiFi Bst DNA polymerase can be used for the classic Sanger manual sequencing and the fluorescent dye-labeled

30   primer sequencing with high processivity and high fidelity.

However, when the protocol was modified for fluorescent dye-labeled ddNTP terminator automated DNA sequencing, numerous "C" peaks and

35   "A" peaks were missing or appeared ambiguous in the sequence plot, especially when the "C" peak was immediately after a "G" peak or after an "A"

36

peak, and when the "A" peak was immediately after
a "G" peak. This selective discrimination against
incorporation of dye-labeled ddCTP and dye-labeled
ddATP is sequence-related since many "C" and "A"
5   peaks following an immediate "C" peak or an
immeciate "T" peak remained strong and correctly
resolved in the same color plot of sequence
analysis. (See Example 9) This phenomenon of
selective discrimination against incorporation of
10 fluorescent dye-labeled ddCTP and dye-labeled
ddATP is observed with all DNA polymerases
isolated from different strains of *Bacillus
stearothermophilus* and *Bacillus caldotenax*, and
appears to be characteristic of DNA polymerases of
15 the mesophilic bacilli.

Example 2:   Mutation of the gene for naturally-
occurring Bst DNA polymerase having proofreading
3'-5' exonuclease activity
        The DNA fragment LF containing the gene
20 initially isolated from the wild Bst 320 has the
following sequence (see SEQ ID NO:1):

DNA sequence (isolated/purified):
    GCCGAAGGGG   AGAAACCGCT   TGAGGAGATG   GAGTTTGCCA
    TCGTTGACGT   CATTACCGAA   GAGATGCTTG   CCGACAAGGC
25 AGCGCTTGTC   GTTGAGGTGA   TGGAAGAAAA   CTACCACGAT
    GCCCCGATTG   TCGGAATCGC   ACTAGTGAAC   GAGCATGGGC
    GATTTTTTAT   GCGCCCGGAG   ACCGCGCTGG   CTGATTCGCA
    ATTTTTAGCA   TGGCTTGCCG   ATGAAACGAA   GAAAAAAAGC
    ATGTTTGACG   CCAAGCGGGC   AGTCGTTGCC   TTAAAGTGGA
30 AAGGAATTGA   GCTTCGCGGC   GTCGCCTTTG   ATTTATTGCT
    CGCTGCCTAT   TTGCTCAATC   CGGCTCAAGA   TGCCGGCGAT
    ATCGCTGCGG   TGGCGAAAAT   GAAACAATAT   GAAGCGGTGC
    GGTCGGATGA   AGCGGTCTAT   GGCAAAGGCG   TCAAGCGGTC
    GCTGCCGGAC   GAACAGACGC   TTGCTGAGCA   TCTCGTTCGC

```
     AAAGCGGCAG  CCATTTGGGC  GCTTGAGCAG  CCGTTTATGG
     ACGATTTGCG  GAACAACGAA  CAAGATCAAT  TATTAACGAA
     GCTTGAGCAC  GCGCTGGCGG  CGATTTTGGC  TGAAATGGAA
     TTCACTGGGG  TGAACGTGGA  TACAAAGCGG  CTTGAACAGA
  5  TGGGTTCGGA  GCTCGCCGAA  CAACTGCGTG  CCATCGAGCA
     GCGCATTTAC  GAGCTAGCCG  GCCAAGAGTT  CAACATTAAC
     TCACCAAAAC  AGCTCGGAGT  CATTTTATTT  GAAAAGCTGC
     AGCTACCGGT  GCTGAAGAAG  ACGAAAACAG  GCTATTCGAC
     TTCGGCTGAT  GTGCTTGAGA  AGCTTGCGCC  GCATCATGAA
 10  ATCGTCGAAA  ACATTTTGCA  TTACCGCCAG  CTTGGCAAAC
     TGCAATCAAC  GTATATTGAA  GGATTGTTGA  AAGTTGTGCG
     CCCTGATACC  GGCAAAGTGC  ATACGATGTT  CAACCAAGCG
     CTGACGCAAA  CTGGGCGGCT  CAGCTCGGCC  GAGCCGAACT
     TGCAAAACAT  TCCGATTCGG  CTCGAAGAGG  GGCGGAAAAT
 15  CCGCCAAGCG  TTCGTCCCGT  CAGAGCCGGA  CTGGCTCATT
     TTCGCCGCCG  ATTACTCACA  AATTGAATTG  CGCGTCCTCG
     CCCATATCGC  CGATGACGAC  AATCTAATTG  AAGCGTTCCA
     ACGCGATTTG  GATATTCACA  CAAAAACGGC  GATGGACATT
     TTCCAGTTGA  GCGAAGAGGA  AGTCACGGCC  AACATGCGCC
 20  GCCAGGCAAA  GGCCGTTAAC  TTCGGTATCG  TTTACGGAAT
     TAGCGATTAC  GGATTGGCGC  AAAACTTGAA  CATTACGCGC
     AAAGAAGCTG  CCGAATTTAT  CGAACGTTAC  TTCGCCAGCT
     TTCCGGGCGT  AAAGCAGTAT  ATGGAAAACA  TAGTGCAAGA
     AGCGAAACAG  AAAGGATATG  TGACAACGCT  GTTGCATCGG
 25  CGCCGCTATT  TGCCTGATAT  TACAAGCCGC  AATTTCAACG
     TCCGCAGTTT  TGCAGAGCGG  ACGGCCATGA  ACACGCCAAT
     TCAAGGAAGC  GCCGCTGACA  TTATTAAAAA  AGCGATGATT
     GATTTAGCGG  CACGGCTGAA  AGAAGAGCAG  CTTCAGGCTC
     GTCTTTTGCT  GCAAGTGCAT  GACGAGCTCA  TTTTGGAAGC
 30  GCCAAAAGAG  GAAATTGAGC  GATTATGTGA  GCTTGTTCCG
     GAAGTGATGG  AGCAGGCCGT  TACGCTCCGC  GTGCCGCTGA
     AAGTCGACTA  CCATTACGGC  CCAACATGGT  ATGATGCCAA
     ATAA    (1764 nucleotides total)
```

Site directed mutagenesis was performed
35   as described by Kunkel et al.(14) The DNA fragment
(designated "LF") containing the gene for Bst DNA

polymerase having proofreading exonuclease
activity was cloned from the expression vector
pYZ23/LF into plasmid pUC119. The constructed
plasmid pUC119/LF was then transformed into E.

5   coli CJ236, a mutant of E. coli that lacks the
enzymes dUTPase and uracil N-glycosylase.
Therefore, when grown in a medium supplemented
with uridine, this mutant of E. coli as well as
the plasmids in the cells will incorporate

10   deoxyuridine into the DNA in place of thymidine
and the uracils will not be removed readily.

As the constructed plasmid grew in the
cells of E. coli CJ236 and in the presence of
uracil and M13KO7 helper phage, some of the normal

15   thymidine bases of the DNA in the newly produced
single-stranded pUC119/LF were replaced by
uracils. These uracil-containing DNAs were used
as the template in vitro for the production of a
complementary oligonucleotide that contained the

20   desired DNA sequence alteration, but with only
dTMPs and not dUMP residues.

In practice, the expression vector
pYZ23/LF was digested with restriction enzymes Eco
RI and Bam HI, and the DNA fragment LF was

25   separated and cloned into plasmid pUC119 which had
been previously digested with the same restriction
enzymes. The constructed plasmid pUC119/LF was
then transformed into E. coli CJ236. For gaining
the uracil-containing single-stranded pUC119/LF, a

30   colony of E. coli CJ236 containing pUC119/LF was
selected and inoculated into 2 ml of 2 x YT medium
which was supplemented with 0.25 ug/ml of uridine
and 2×10^8 to 4×10^8 pfu/ml of M13KO7 as helper
phage. After incubation at 37°C with strong

35   agitation for 1 hour, a kanamycin solution (25
mg/ml in H_2O) was added to the culture to a final
concentration of 70 ug/ml. The incubation was

39

allowed to continue for another 14-18 hours at
37°C with strong agitation. Then 1.5 ml of the
infected culture was transferred to a
microcentrifuge tube, and centrifuged at 12,000 x
5   g for 5 minutes at 4°C. The uracil-containing
single-stranded pUC119/LF was precipitated and
purified from the supernatant according to
standard PEG/NaCl and ethanol procedures.

        After performing a series of
10   experiments, the inventors found that the combined
effects of changing the amino acids leucine-
glutamate-glutamate (LEE) at the location 342-344,
to respectively threonine-proline-leucine (TPL),
and the amino acid phenylalanine (F) at location
15   422, to tyrosine (Y) in the peptide structure of
HiFi Bst DNA polymerase markedly reduced its
selective discrimination against incorporation of
fluorescent dye-labeled ddCTP and dye-labeled
ddATP to such a level that direct automated
20   fluorescent DNA sequence (although not cycle-
sequencing) can be performed with the dye-
terminator technology when the mutated enzyme of
the current invention is used. It is of interest
to note that this modified HiFi Bst, now referred
25   to as HiFi Bst-II DNA polymerase, exhibits the
function of preferentially incorporating more
fluorescent dye-labeled ddCTP and dye-labeled
ddATP onto the 3' end dGMP and the dAMP bases of
the extending DNA strands during enzymatic
30   reaction, than the unmodified naturally occurring
HiFi Bst polymerase.

        The end result is the recovery of the
"C" and "A" peaks which otherwise would have been
missing or ambiguous on the sequence analysis
35   color plot. At the same time, the modified enzyme
did not indiscriminately generate an excess amount
of dye-labeled "G" terminated or dye-labeled "T"

terminated DNA fragments. Even the "C" and "A" peaks were not uniformly raised in a blanket manner, but only raised in the formerly depressed locations after a "G" and/or an "A". (See Example 9). Thus, this genetic modification of the HiFi Bst to HiFi Bst-II results in a DNA polymerase that reduces the selective discrimination against incorporation of the fluorescent dye-labeled ddCTP and dye-labeled ddATP, rather than merely increases the ability of the parent enzyme to incorporate these dye-labeled dideoxynucleotides.

HiFi Bst-II, and the other novel similar DNA polymerases of this invention, can be used for the classic radiolabeling Sanger method. (See Example 8.) HiFi Bst-II appears to generate a better sequencing pattern than HiFi Bst and requires less ddNTPs to terminate the extending reaction (Figure 3). For instance, in the optimized reaction mixture for the unmodified HiFi Bst DNA polymerase, the ddNTP/dNTP ratios in the A, C, G and T mix were 40, 6.25, 18.25 and 18.72, respectively. In the optimized reaction mixture for the modified HiFi Bst-II DNA polymerase, the corresponding ddNTP/dNTP ratios in the A, C, G and T mix were 40, 2.5, 6.25 and 6.25, respectively. Therefore, there was an up-to about three-fold reduction in the amount of ddNTPs used after genetic modification of the naturally-occurring DNA polymerase.

For the radiolabeling classic Sanger method of DNA sequencing, the optimized reaction mixtures for either HiFi Bst or HiFi Bst-II must contain much more ddNTPs than dNTPs to generate a ladder of DNA fragments for sequencing analysis because the DNA polymerases of the mesophilic bacilli tend to incorporate dNTPs more efficiently than ddNTPs. The above-described genetic

41

modification appears to increase the ability of
the naturally-occurring enzymes to incorporate
ddNTP in the presence of a corresponding competing
dNTP to about three-fold at the concentration

5   ratios commonly used for DNA sequencing. However,
if much higher concentrations of the nucleotides
were used for the experiment, and the ddNTP/dNTP
ratio was reduced to a level that is suboptimal
for DNA sequencing (for instance at a ratio of

10   1/3), the increased ability for incorporating
ddNTPs after modification of the enzyme could be
dramatized. (See Example 6, Figure 2).

        Similar to the results obtained with
radiolabeling Sanger method, both HiFi Bst and

15   HiFi Bst-II can be adapted for fluorescent dye-
labeled primer automated DNA sequencing and
produce comparable results without selective
suppression of any specific fluorescent peaks in
the sequencing plot (see Example 9) although the

20   peaks generated by HiFi Bst-II appear to be more
even than those by HiFi Bst.

        To change amino acids leucine, glutamic
acid and glutamic acid (LEE) at positions 342-344,
respectively in the Bst polymerase into threonine,

25   proline and leucine (TPL), respectively, Primer 1
was designed as following (see SEQ ID NO 10):
        5'-CATTCCGATTCGGACCCCACTGGGGCGGAAAATCCG-3

        To change amino acid phenylalanine (F)
at position 422 in the Bst DNA polymerase into

30   tyrosine (Y), Primer 2 was designed as following
(see SEQ ID NO: 9):
        5'-GCCGTTAACTACGGTATCGTTTACGG-3'

        After phosphorylation of the 5' ends of
the oligonucleotides by T4 polynucleotide kinase,

35   the two primers designed above were annealed to
the single-stranded uracil-containing pUC119/LF
purified from above. In the presence of the usual

42

dNTPs (dATP, dCTP, dGTP and dTTP), T4 DNA
polymerase was used to synthesize _in vitro_ the
strands of DNA complementary to the uracil-
containing pUC119/LF template, and T4 ligase was
5   used to ligate the synthesized strands to form a
complete double-stranded plasmid which was
composed of one single-stranded, not mutagenic,
uracil-containing pUC119/LF and one complementary
single-stranded, mutagenic, thymidine-containing
10   DNA fragment that had been altered by primer 1 and
primer 2 described above. These newly formed
double-stranded plasmids were then transformed
into E. coli JM109. The template strand was
rendered biologically inactive. The transformed
15   strain of E. coli JM109 whose plasmids contained
the mutated DNA, now referred to as pUC119/LF-M,
was screened out with DNA sequencing of its
plasmids.

Example 3: Cloning and expression of the modified
20   Bst DNA polymerase having both ability to reduce
selective dye-labeled ddNTP discrimination and
proofreading 3'-5' exonuclease activity

The plasmid pUC119/LF-M was prepared
from the strain of _Escherichia_ _coli_ JM109
25   containing the mutated DNA. The mutated DNA
fragment (LF-M) containing the mutated gene for
the Bst polymerase was recombined back into the
expression vector pYZ23. The constructed plasmid
pYZ23/LF-M was then transformed into _Escherichia_
30   _coli_ JF1125. The mutation was further confirmed by
double-stranded dideoxy DNA sequencing of isolated
plasmid.
The strain of _Escherichia_ _coli_ JF1125
containing pYZ23/LF-M was inoculated into LB
35   culture containing 100µg/ml ampicillin, and was

43

incubated overnight at 30°C. The overnight culture
was inoculated into a large volume of fresh
culture, and was incubated at 30°C until the OD_{600}
of the culture reached 0.7. The culture was then
5   heated at 41°C for 3 hours for induction. The SDS-
PAGE analysis of the cell extract showed that the
cloned mutated gene for the modified Bst DNA
polymerase was overexpressed.

Example 4:   Isolation and purification of the
10   modified Bst DNA polymerase having both ability to
reduce selective dye-labeled ddNTP discrimination
and proofreading 3'-5' exonuclease activity

The expressed cells of Escherichia coli
JF1125 containing pYZ23/LF-M grown in condition as
15   described above were thawed and washed with buffer
[10mM Tris-HCl(pH7.5 at room temperature), 10mM β-
Mercaptoethanol, 2 mM EDTA, 0.9% NaCl]. The
pellets were then suspended in buffer[50mM Tris-
HCl (pH7.5 at room temperature), 10mM β-
20   Mercaptoethanol, 2mM EDTA, 100μg/ml Lysozyme,
23μg/ml PMSF](4ml/g pellet). After 20 min at room
temperature, the mixture was cooled on salt-ice
and sonicated briefly to complete lysis. The cell
extract obtained by centrifugation at 18,000rpm at
25   4°C for 20 minutes, was then treated step by step
as follows:
(A) The cell extract was heated at 60°C for 30
minutes, and cooled to 4°C, then centrifuged at
15,000rpm at 4°C for 20 minutes;
30   (B) 5% Polymin P was added into supernatant to
0.6%, and mixed quickly for 30 minutes, then
centrifuged;
(C) The pellet was resuspended in Buffer A[50mM
Tris-HCl(pH7.5 at room temperature), 1mM EDTA, 1mM
35   β-Mercaptoethanol] containing 800mM NaCl and 5%

44

Glycerol at 4°C, and then centrifuged;

(D) Ammonium sulfate was added into the
supernatant to 60% saturation at 4°C, and
mixed for 30 minutes, then centrifuged;

5   (E) The ammonium sulfate pellet was resuspended in
30ml of 60% saturated ammonium sulfate at 4°C, and
then recentrifuged;

(F) The pellet was suspended in Buffer A
containing 100mM KCl and dialysed against the same

10   buffer for hours at 4°C, then centrifuged. The
insoluble protein was discard;

(G) The supernatant was added to pass through a
DE-52 column. The column was washed, and the peak
DNA polymerase was eluted using a 100-600 mM KCl

15   linear gradient in Buffer A, concentrated in
Buffer A containing 50%(w/v) PEG-6000, dialyzed in
Buffer A containing 100 mM KCl ;

(H) The solution was then applied to Heparin-
Sepharose CL-4B column. The peak DNA polymerase

20   was eluted with a linear gradient of 100-800 mM
KCl in Buffer A, concentrated and finally dialyzed
in buffer A containing 50% glycerol.

The resulting modified Bst DNA polymerase
has been proved to be homogenous by polyacrylamide

25   gel electrophoresis. And the enzyme obtained was
stored in -20°C.

Example 5: Determination of the thermostability
of unmodified Bst DNA polymerase and modified Bst
DNA polymerase

30   The DNA polymerases of Examples 1 and 5
were incubated at 65°C for 0, 5, 10, 20, 30, 40,

45

50 minutes respectively, and placed into ice-water immediately. The polymerase activity of these DNA polymerases was determined at 60°C.

The polymerase activity of DNA polymerase was determined as follows:

5 x Reaction Solution:

1M Tris-HCl (pH7.6)	16.75 ml
1M MgCl$_2$	1.675 ml
1M β-Mercaptoethanol	0.25 ml
ddH$_2$O	adjusted to 50 ml

Reaction Storage:

5 x Reaction Solution	60μl
dNTPs (1mM each)	10μl
1.5μg/μl DNase I activated	
calf thymus DNA	10μl
ddH$_2$O	10μl
α-^{32}P-dATP	appropriate amt.

Reaction Mixture:

Reaction Storage	30μl
Sample	5μl
ddH$_2$O	65μl

The reaction mixtures were prepared as per the recipe above, and incubated at 60°C for 30 minutes. Then the reaction mixtures were pipetted onto DE-81 filters respectively. After all of the fluid has evaporated, the amount of radioactivity on each filter was measured with scintillation (X_1). The filters were washed three times with 0.3M Na$_2$HPO$_4$ at room temperature, 10 minutes each times, dried at room temperature and then the amount of radioactivity on each filter was measured again (X_2).

The polymerase activity of sample (u/ml) =

46

$$\left[\frac{X_2}{X_1} - \frac{X_{20}}{X_{10}}\right] \times 266$$

(X_{10} and X_{20} are the amount of radioactivity measured with water as control sample)

5   Unit definition of polymerase activity:
One unit is the amount of DNA polymerase required
to incorporate 10 nanomoles of dNTPs into DNA in
30 minutes at 60°C .

            The thermostability of DNA polymerase is
10   expressed with the half life of polymerase
activity at 65°C. Figure 1 shows the comparison of
thermostabilty of HiFi Bst and HiFi Bst-II. The
half life of HiFi Bst at 65°C was 8.5 minutes, and
that of HiFi Bst-II was 16 minutes. HiFi Bst-II
15   was more thermostable than HiFi Bst.

Example 6: Demonstration of increased ddNTP
incorporation by modified Bst DNA polymerase in
suboptimal sequencing conditions

            The following procedure was followed:

20   1. The -20M13 forward primer was
radiolabelled using γ-^{32}P-ATP and T4 Polynucleotide
kinase;
            2. The following components were combined in
a microcentrifuge tube:
25   5×Reaction Buffer   2.0μl
radiolabeled primer   1.0μl (2.5ng)
Template   7.0μl (1μg M13mp18 ssDNA)
The final volume was 10μl. The contents were mixed
and spun for 2-3 seconds;

30   3. The tube were placed in a 75°C water bath
for 5 minutes. Then the tube was allowed to cool
slowly to ambient temperature over a course of 10

47

minutes;

    4. 1.0μl of modified Bst DNA polymerase (of
Example 5) (1u/μl) was added. The mixture was
mixed gently and spun for 2-3 seconds;

5   5. 4 tubes were labelled "A", "C", "G", "T",
respectively and 2μl of each premixed nucleotide
solution and 2.5μl of main mixture (from step 3)
was added to the respective reaction
tube;

10   6. The tubes were incubated at 65°C for 15
minutes;

    7. The reactions were stopped by adding 4.0μl
of Stop Solution(95% deionized formamide, 10mM
EDTA, 0.05% xylene cyanol FF, 0.05% bromophenol
15   blue) to each tube;

    8. The samples were denatured at 90°C for 2
minutes, and immediately placed on ice;

    9. 4-5μl of samples were loaded onto each
lane of 6% (8M urea) sequencing gel, and
20   electrophoresis was carried out.

Note: The Components of the Premixed Nucleotide
Solutions:
        A mix: dNTPs 120μM, ddATP 40μM
        C mix: dNTPs 120μM, ddCTP 40μM
25   G mix: dNTPs 120μM, ddGTP 40μM
        T mix: dNTPs 120μM, ddTTP 40μM

        Figure 2 shows the comparison of ddNTP
incorporation of HiFi Bst-II DNA polymerase and
HiFi Bst DNA polymerase. In this radiolabeling

48

DNA sequencing experiment, high concentrations of
nucleotides were used in the reaction mixture and
the ddNTP/dNTP ratio was reduced to a level (1/3)
that is lower than the optimal range for DNA

5   sequencing. HiFi Bst-II is shown to have more
effective ddNTP incorporation. The DNA synthesis
was often terminated by ddNTP incorporation in the
HiFi Bst-II mixture, and the result showed uniform
bands with synthesized small or large DNA

10   fragments. As a contrast, HiFi Bst had a lower
ddNTP incorporation. The DNA synthesis by HiFi Bst
was less terminated, and most of the synthesized
products were the larger DNA fragments.

Example 7:   Preparation of Denatured Double-
15   stranded DNA Template:

The following procedure was carried out.

1. Double-stranded DNA(about 3-5μg) was
adjusted to a final volume of 10μl with TE (10mM
Tris-HCl, 1mM EDTA, pH8.0);

20   2. 10μl of 0.4N NaOH, 0.4mM EDTA, was added;

3. The mixture was incubated at 65°C for 15
minutes;

4. 2μl of 2M sodium acetate, pH4.5, and 55μl
cold ethanol was added, and the mixture was placed
25   in ice-water bath for 5 minutes;

5. The mixture was spun in a microcentrifuge
at 4°C, 12500rpm for 5 minutes;

6. The supernatant was drawn off and the
pellet was washed with 200μl of 70% ethanol;

49

7. The pellet was dried under vacuum for 2-3
minutes, and the DNA was dissolved in appropriate
solution.

Example 8:   DNA sequencing using unmodified Bst
5   DNA polymerase/modified Bst DNA polymerase with
radiolabeled dATP for single- or denatured double-
stranded DNA template
          The following procedure was carried out.
       1. The following components were combined in
10   a labeled microcentrifuge tube:

5×Reaction Buffer   2.0μl
Primer   1.0μl (2.5-5.0ng)
Template   7.0μl (250-500ng ss DNA
                           or 1-3μg denatured ds DNA)
15   The final volume was 10μl. The contents were mixed
and spun for 2-3 seconds;

       2. The tube were placed in a 75°C water bath
for 5 minutes. Then the tubes were allowed to cool
slowly to ambient temperature over a course of 10
20   minutes;
(Note: Step 2 is optional for single-stranded
template, and may be omitted at appropriate.)

       3. 1.0μl of HiFi Bst/HiFi Bst-II (1u/μl) and
1.0μl of [α-^{32}P]dATP was added, and the mixture was
25   mixed gently and spun for 2-3 seconds;

       4. 4 tubes "A", "C", "G", "T" were labelled,
and 2μl of each premixed nucleotide solution and
2.5μl of main mixture (from step 3) was added to
the respective reaction tube;

30   5. The tubes were incubated at 65°C for 2
minutes;

50

6. 2.0µl of 0.5mM dNTPs was added to each
tube, and the tubes were mixed gently, spun for 2-
3 seconds, and incubated at 65°C for 2 minutes;

7. The reactions were stopped by adding 4.0µl
5   of Stop Solution(95% deionized formamide, 10mM
EDTA, 0.05% xylene cyanol FF, 0.05% bromophenol
blue) to each tube;

8. The samples were denatured at 90°C for 2
minutes, and immediately placed on ice;

10   9. 2-3µl of the samples were loaded onto each
lane of 6% (8M urea) sequencing gel, and
electrophoresis was carried out.

Note: The Components of the Premixed Nucleotide
Solutions for HiFi Bst:
15   A mix: dATP0.62µM, dCTP 62µM, dGTP 62µM, dTTP
62µM, ddATP 25µM;
C mix: dATP 0.8µM, dCTP 8µM, dGTP 80µM, dTTP
80µM, ddCTP 50µM;
G mix: dATP 0.8µM, dCTP 80µM, dGTP 4µM, dTTP
20   80µM, ddGTP 75µM;
T mix: dATP 0.8µM, dCTP 80µM, dGTP 80µM, dTTP
8µM, ddTTP 150µM.

The Components of the Premixed Nucleotide
Solutions for HiFi Bst-II:
25   A mix: dATP 0.8µM, dCTP 80µM, dGTP 80µM, dTTP
80µM, ddATP 25µM;
C mix: dATP 0.8µM, dCTP 8µM, dGTP 80µM, dTTP
80µM, ddCTP 20µM;
G mix: dATP 0.8µM, dCTP 80µM, dGTP 8µM, dTTP
30   80µM, ddGTP 50µM;
T mix: dATP 0.8µM, dCTP 80µM, dGTP 80µM, dTTP
8µM, ddTTP 50µM.

Figure 3 shows the comparison of
radiolabeled DNA sequencing with HiFi Bst and HiFi
Bst-II in their respective optimized reaction
mixtures. The bands on a DNA sequencing gel with

5   HiFi Bst-II were more uniform than those with HiFi
Bst. The sequence pattern using HiFi Bst-II was
better than that using HiFi Bst. This made the gel
with HiFi Bst-II even easier to read. Moreover,
the ratio of dideoxy-nucleotide concentration to

10   deoxy-nucleotide concentration in Premixed
Nucleotide Solutions for HiFi Bst-II was lower
than that for HiFi Bst. For instance, the
ddATP/dATP, ddCTP/dCTP, ddGTP/dGTP and ddTTP/dTTP
ratios were 40, 6.25, 18.75 and 18.75,

15   respectively, for HiFi Bst. The corresponding
ratios for HiFi Bst-II were 40, 2.5, 6.25 and
6.25, respectively. Therefore, the concentrations
of ddNTPs used in the termination reaction are
reduced to about three-folds after HiFi Bst has

20   been genetically modified.

Example 9:   DNA sequencing using unmodified Bst
DNA polymerase/modified Bst DNA polymerase with
dye-primers for single- or denatured double-
stranded DNA template

25   The following procedure was carried out.

1. The following was combined: 4.0μl DNA
template (300-600ng for ssDNA or 1-2μg denatured
ds DNA) with 5.0μl 5xReaction Buffer. The mixture
was mixed and spun for 2-3 seconds in a

30   microcentrifuge;

2. 4 tubes were labeled "A", "C", "G", "T"
and the pre-mixed dATP, dCTP, dGTP, dTTP reagents
were added to each reaction tube:

                        Reaction tube

52

	A	C	G	T
A-REG primer (0.2uM)	1.0ul			
C-FAM primer (0.2uM)		1.0ul		
G-TMR primer (0.4uM)			1.0ul	
T-ROX primer (0.4uM)				1.0ul
A terminator mix	2.0ul			
C terminator mix		2.0ul		
G terminator mix			2.0ul	
T terminator mix				2.0ul
5 x Reaction buffer				
with DNA template	2.0ul	2.0ul	2.0ul	2.0ul
Total volume	5.0ul	5.0ul	5.0ul	5.0ul

3. The tubes were placed in 75°C bath for 5
minutes, and allowed to cool slowly to ambient
temperature over the course of 10 minutes;
(Note: Step 3 is optional for single-stranded
template, and may be omitted as appropriate.)

4. 1μl of HiFi Bst/HiFi Bst-II (0.5u/μl) was
added to each tube, and the tubes were spun for 2-
3 seconds;

5. The tubes were incubated at 65°C for 5
minutes;

6. The contents of the "A", "C", "G" and "T"
tubes were pooled, and 1.5μl of 7.5M ammonium
acetate and 55μl of ethanol was added. The
mixtured was mixed in a vortex and then placed on
ice for 20 minutes;

7. The mixture was centrifuged at 12,500rpm
for 20 minutes at 4°C;

8. The supernatant was drawn off, and the

53

pellet was washed with 200µl of 70% ethanol;

9. The pellet was vacuum dried for 2-3
minutes, and resuspended in 4µl of loading buffer
(5:1 deionized formamide : 25mM EDTA with 50mg/ml
5   Blue Dextran);

10. The sample was heated at 75°C for 2-3
minutes, and immediately placed on ice;

11. 2-3 µl of sample was loaded onto a lane
of the 4% (6M urea) sequencing gel, and ABI PRISM™
10   377 DNA Sequencer (from Perkin Elmer) was used to
collect data.

Note: Dye primer: DYEnamic Energy Transfer Dye
Primers (from Amersham):
-21 M13 forward: 5'-FAM-S^{T}SSSSSTGT·AAAACGACGGCCAGT-
15   3' (SEQ ID NO:11)
TS=1'2'-dideoxyribose
T·=T attached with Dye 2(A-REG, C-FAM, G-TMR, T-
ROX)

Figure 4 and Figure 5 show the results
20   of dye-primer DNA sequencing with HiFi Bst and
HiFi Bst-II. Both DNA polymerases generated
similar sequencing results although the peaks on
the color plot by HiFi Bst II appear to be more
even in height.

25   Example 10: DNA sequencing using unmodified Bst
DNA polymerase/modified Bst DNA polymerase with
dye-terminators for single- or denatured double-
stranded DNA template
The following procedure was carried out.

30   1. The following components were combined in
a labeled microcentrifuge tube:

54

```
                    5×Reaction Buffer    4.0µl
                    Template             8.0µl  (2-3µg ss DNA
                                         or  4-6µg denatured
          ds DNA)
5                   Primer               2.0µl  (5-10ng)
```

The final volume was 14µl. The contents were
mixed and spun for 2-3 seconds;

2. The tube was placed in a 75°C water bath
for 5 minutes;

10   3. The tube was allowed to cool slowly to
ambient temperature over a course of 10 minutes;
(Note: Steps 2 and 3 are optional for single-
stranded template, and may be omitted as
appropriate.)

15   4. 1.0µl of HiFi Bst/HiFi Bst-II (1-2u/µl),
5µl of nucleotides premix (containing Perkin
Elmer-ABI fluorescent dye-labeled nucleotide
terminators), were added and the tube was spun for
2-3 seconds;

20   5. The mixture was incubated at 65°C for 10
minutes;

6. 80µl of H_2O was added to the reaction mix,
and the dye terminators were extracted with 100µl
of phenol:H_2O:chloroform (68:18:14) reagent twice.
25   The sample was vortexed and centrifuged, and the
aqueous upper layer was transferred to a clean
tube;

7. To the tube was added 15µl of 2M sodium
acetate, pH 4.5, and 300µl of ethanol, and the
30   tube was vortexed and placed in ice-water bath for
20 minutes;

                              55

8. The tube was centrifuged with 12,500rpm
for 20 minutes at 4°C;

.9. The supernatant was drawn off, and the
pellet was washed with 200µl of 70% ethanol;

5   10. The pellet was vacuum dried for 2-3
minutes, and resuspended in 4µl of loading buffer
(5:1 deionized formamide : 25mM EDTA with 50mg/ml
Blue Dextran);

11. The sample was heated at 90°C for 2-3
10   minutes, and immediately placed on ice;

12. 2-3µl of sample was loaded onto a lane of
the 4% (6M urea) sequencing gel, and ABI PRISM™
377 DNA Sequencer (from Perkin Elmer) was employed
to collect data, using appropriate amounts of
15   nucleotide pre-mixed reagents.

Figure 6 and Figure 7 show the results of
dye-terminator DNA sequencing with HiFi Bst and
HiFi Bst-II. There was data lost in dye-terminator
DNA sequencing with HiFi Bst, especially the "C"
20   after "G" or "A" and "A" after "G". In Figure 8,
corrections of the missing or ambiguous bases,
according to the known pGEM sequence, have been
indicated below the letters "N" or below the
incorrect base letters. This problem caused
25   ambiguity in DNA sequencing. But it was resolved
in dye-terminator DNA sequencing with the modified
Bst DNA polymerase of this invention.

REFERENCES
1. Sanger, F., Nicklen, S. & Coulson, A.R. Proc.
30   Nat. Acad.Sci., USA 74: 5463-5467. 1977.
2.   CURRENT PROTOCOLS IN MOLECULAR BIOLOGY,

56

Ausubel, F.M. et. Al. (Editors) Vol. I., John
Wiley & Sons, Inc. 1995. pp 7.4.17-7.4.24.

3. Ibid p. 7.4.31.

4. Ye, S.Y. & Hong, G.F., Scientia Sinica (Series
B) 30: 503-506. 1987.

5. In Ref. 2, p. 7.4.18 Table 7.4.2.

6. EPICENTRE TECHNOLOGIES CATALOG,1994/95 Products
for Molecular & Cellular Biology, Page 1, " What's
new in this catalog?"

7. Okazaki, T. & Kornberg, A. J. Biol. Chem. 239:
259-268. 1964.

8. Jacobsen, H., Klenow, H. & Overgard-Hansen, K.
Eur. J. Biochem. 45: 623-627. 1974.

9. McClary, J., Ye, s.Y., Hong, G.F. & Witney, F.
DNA Sequence 1: 173-180. 1991.

10. Mead, D. A., McClary, J.A., Luckey, J.A., et
Al. BioTechniques 11: 76-87. 1991.

11. Earley, J.J., Kuivaniemi, H. Prockop, D.J. &
Tromp, G. BioTechniques 17: 156-165,1994.

12. Mardis, E.R. & Bruce, A.R. BioTechniques 7:
840-850. 1989.

13. Chissoe, S.L. et al., Methods: A Companion to
Methods in Enzymology, 3, 555-65, 1991

14. Kunkel, T.A. et al., Methods Enzymol. 154:367-
382, 1987

All references mentioned herein are
incorporated in their entirety by reference.

<u>WHAT IS CLAIMED IS</u>:

1. A modified DNA polymerase which
during DNA sequencing effectively incorporates
fluorescent dye-labeled dideoxynucleotide
terminators ddCTP, ddATP, ddTTP and ddGTP, and
their analogs, and reduces selective
discrimination against incorporation of
fluorescent dye-labeled dideoxynucleotide
terminators ddCTP and ddATP,

wherein the DNA polymerase in its unmodified
state selectively discriminates against
incorporation of fluorescent dye-labeled
dideoxynucleotide terminators ddCTP and ddATP but
does not discriminate against incorporation of
fluorescent dye-labeled dideoxynucleotide
terminators ddTTP and ddGTP.

2. The DNA polymerase according to
claim 1 which is a modified *Bacillus
stearothermophilus* DNA polymerase.

3. The DNA polymerase according to
claim 2 which has an amino acid sequence that
shares not less than 95% homology of a DNA
polymerase isolated from a strain of *Bacillus
stearothermophilus*, *Bacillus caldotenax* or
Bacillus caldolyticus.

4. The DNA polymerase according to
claim 1 wherein the DNA polymerase is a modified
DNA polymerase obtained from a mesophilic
bacterium.

5. The DNA polymerase according to
claim 1 which is a thermostable DNA polymerase
having proofreading 3'-5' exonuclease activity
during DNA sequencing of a DNA strand from a

template, such that the DNA polymerase functions
to excise mismatched nucleotides from the 3'
terminus of the DNA strand at a faster rate than
the rate at which the DNA polymerase functions to
remove nucleotides matched correctly with
nucleotides of the template.

6.   A modified *Bacillus*
stearothermophilus DNA polymerase which during DNA
sequencing effectively incorporates fluorescent
dye-labeled dideoxynucleotide terminators ddCTP,
ddATP, ddTTP and ddGTP, and reduces selective
discrimination against incorporation of
fluorescent dye-labeled dideoxynucleotide
terminators ddCTP and ddATP,
        wherein the DNA polymerase in its unmodified
state selectively discriminates against
incorporation of fluorescent dye-labeled
dideoxynucleotide terminators ddCTP and ddATP but
does not discriminate against incorporation of
fluorescent dye-labeled dideoxynucleotide
terminators ddTTP and ddGTP,
        wherein the DNA polymerase has proofreading
3'-5' exonuclease activity during DNA sequencing
of a DNA strand from a template, such that the DNA
polymerase functions to excise mismatched
nucleotides from the 3' terminus of the DNA strand
at a faster rate than the rate at which the DNA
polymerase functions to remove nucleotides matched
correctly with nucleotides of the template.

7.   The DNA polymerase according to
claim 6, which has the amino acid sequence SEQ ID
NO:4.

8.   The DNA polymerase according to
claim 6, which is encoded by a DNA segment having

59

the nucleotide sequence of SEQ ID:NO 3.

9.   A host cell which produces a
modified DNA polymerase which during DNA
sequencing effectively incorporates fluorescent
dye-labeled dideoxynucleotide terminators ddCTP,
ddATP, ddTTP and ddGTP, and reduces selective
discrimination against incorporation of
fluorescent dye-labeled dideoxynucleotide
terminators ddCTP and ddATP,
        wherein the DNA polymerase in its unmodified
state selectively discriminates against
incorporation of fluorescent dye-labeled
dideoxynucleotide terminators ddCTP and ddATP but
does not discriminate against incorporation of
fluorescent dye-labeled dideoxynucleotide
terminators ddTTP and ddGTP.

10.   The host cell according to claim 9,
wherein the modified DNA polymerase has
proofreading 3'-5' exonuclease activity during DNA
sequencing of a DNA strand from a template, such
that the DNA polymerase functions to excise
mismatched nucleotides from the 3' terminus of the
DNA strand at a faster rate than the rate at which
the DNA polymerase functions to remove nucleotides
matched correctly with nucleotides of the
template.

11.   The host cell according to claim 9,
which produces a DNA polymerase having the amino
acid sequence SEQ ID NO:4.

12.   The host cell according to claim 9,
wherein the DNA polymerase is encoded by the
nucleotide sequence SEQ ID NO:3.

13.   The host cell of claim 9, which is
E. coli.

14.   A method of sequencing a DNA strand
comprising the steps of:

i)   hybridizing a primer to a DNA template to
be sequenced;

ii)   extending the primer using the modified
DNA polymerase of claim 1, in the presence of
adequate amounts of nucleotide bases dATP, dGTP,
dCTP and dTTP, or their analogs, and fluorescent
dye-labeled dideoxynucleotide terminators,

under such conditions that the DNA strand is
sequenced.

15.   The method according to claim 14,
wherein the DNA polymerase is a modified *Bacillus
stearothermophilus* DNA polymerase.

16.   The method according to claim 14,
wherein the modified DNA polymerase has an amino
acid sequence that shares not less than 95%
homology of a DNA polymerase isolated from a
strain of *Bacillus stearothermophilus, Bacillus
caldotenax* or *Bacillus caldolyticus.*

17.   The method according to claim 14,
wherein the modified DNA polymerase has the amino
acid sequence SEQ ID NO:4.

18.   The method according to claim 14,
wherein the DNA polymerase is encoded by the
nucleotide sequence SEQ ID NO:3.

19.   The method according to claim 14,
wherein the DNA polymerase is a modified DNA
polymerase obtained from a mesophilic bacterium.

20. The method according to claim 14, wherein the DNA polymerase is a thermostable DNA polymerase having proofreading 3'-5' exonuclease activity during DNA sequencing of a DNA strand from a template, such that the DNA polymerase functions to excise mismatched nucleotides from the 3' terminus of the DNA strand at a faster rate than the rate at which the DNA polymerase functions to remove nucleotides matched correctly with nucleotides of the template.

21. The method according to claim 20, wherein the DNA polymerase is a modified *Bacillus stearothermophilus, Bacillus caldotenax* or *Bacillus caldolyticus* DNA polymerase.

22. The method according to claim 21, wherein the DNA polymerase has the amino acid sequence of SEQ ID NO:4.

23. The method according to claim 21, wherein the DNA polymerase is encoded by a DNA segment having the nucleotide sequence of SEQ ID:NO 3.

24. A method for producing a modified form of a DNA polymerase which during DNA sequencing selectively discriminates against incorporation of fluorescent dye-labeled dideoxynucleotide terminators ddCTP and ddATP but does not discriminate against incorporation of fluorescent dye-labeled dideoxynucleotide terminators ddTTP and ddGTP, comprising the step of

modifying a DNA polymerase which has an amino acid sequence that shares not less than 95% homology of a DNA polymerase isolated from a

strain of *Bacillus stearothermophilus*, *Bacillus
caldotenax* or *Bacillus caldolyticus*, so that the
modified DNA polymerase includes threonine,
proline and leucine at positions 342-344,
respectively, and tyrosine at position 422.

25.   The method according to claim 24,
wherein the DNA polymerase has proofreading 3'-5'
exonuclease activity during DNA sequencing of a
DNA strand from a template, such that the DNA
polymerase functions to excise mismatched
nucleotides from the 3' terminus of the DNA strand
at a faster rate than the rate at which the DNA
polymerase functions to remove nucleotides matched
correctly with nucleotides of the template.

26.   The method according to claim 24,
wherein the DNA polymerase has the amino acid
sequence of SEQ ID:NO 4.

27.   The method according to claim 24,
wherein the DNA polymerase is encoded by a DNA
segment having the nucleotide sequence of SEQ
ID:NO 3.

28.   A method for producing a modified
form of a DNA polymerase which during DNA
sequencing selectively discriminates against
incorporation of fluorescent dye-labeled
dideoxynucleotide terminators ddCTP and ddATP but
does not discriminate against incorporation of
fluorescent dye-labeled dideoxynucleotide
terminators ddTTP and ddGTP, comprising the step
of
      modifying a nucleotide sequence encoding a
DNA polymerase which has an amino acid sequence
that shares not less than 95% homology of a DNA

polymerase isolated from a strain of *Bacillus stearothermophilus* or *Bacillus caldotenax*, so that the modified nucleotide sequence encodes threonine, proline and leucine residues at positions 342-344, respectively, and a tyrosine residue at position 422.

29.   A DNA construct comprising:

(i) a nucleotide sequence encoding a DNA polymerase which has an amino acid sequence that shares not less than 95% homology of a DNA polymerase isolated from a strain of *Bacillus stearothermophilus* or *Bacillus caldotenax*, which nucleotide sequence encodes threonine, proline and leucine residues at positions 342-344, respectively, and a tyrosine residue at position 422; and

(ii) a vector, for introducing the DNA construct into euaryotic or procaryotic host cells.

30.   The DNA construct according to claim 29 wherein the vector is a cloning vector or an expression vector.

31.   A kit for direct DNA sequencing comprising the modified DNA polymerase of claim 1 and at least one fluorescent dye-labeled ddNTP.

32.   The kit according to claim 31 wherein the modified DNA polymerase is a modified *Bacillus stearothermophilus*, *Bacillus caldotenax* or *Bacillus caldolyticus* DNA polymerase.

33.   The kit according to claim 31 wherein the modified DNA polymerase has an amino acid sequence that shares not less than 95%

64

homology of a DNA polymerase isolated from a
strain of *Bacillus stearothermophilus*, *Bacillus
caldotenax* or *Bacillus caldolyticus*.

34.   The kit according to claim 31
wherein the modified DNA polymerase is modified
DNA polymerase obtained from a mesophilic
bacterium.

35.   The kit according to claim 31
wherein the modified DNA polymerase has
proofreading 3'-5' exonuclease activity during DNA
sequencing of a DNA strand from a template, such
that the DNA polymerase functions to excise
mismatched nucleotides from the 3' terminus of the
DNA strand at a faster rate than the rate at which
the DNA polymerase functions to remove nucleotides
matched correctly with nucleotides of the
template.

36.   The kit according to claim 31,
wherein the modified DNA polymerase is modified
Bacillus stearothermophilus DNA polymerase which
has proofreading 3'-5' exonuclease activity during
DNA sequencing of a DNA strand from a template,
such that the DNA polymerase functions to excise
mismatched nucleotides from the 3' terminus of the
DNA strand at a faster rate than the rate at which
the DNA polymerase functions to remove nucleotides
matched correctly with nucleotides of the
template.

37.   The kit according to claim 31
wherein the modified DNA polymerase has the amino
acid sequence SEQ ID NO:4.

38.   The kit according to claim 31

65

wherein the modified DNA polymerase is encoded by a DNA segment having the nucleotide sequence of SEQ ID:NO 3.

39.   A modified nucleotide sequence encoding a DNA polymerase which has an amino acid sequence that shares not less than 95% homology of a DNA polymerase isolated from a strain of *Bacillus stearothermophilus* or *Bacillus caldotenax*, which nucleotide sequence encodes threonine, proline and leucine residues at positions 342-344, respectively, and a tyrosine residue at position 422.

40.   The modified nucleotide sequence according to claim 39 which has the nucleotide sequence of SEQ ID:NO 3.

41.   A method of dye-labeled primer sequencing of a DNA strand comprising the steps of:

    i)   hybridizing a fluorescent dye-labeled primer to a DNA template to be sequenced;

    ii)   extending the primer using the modified DNA polymerase of claim 1, in the presence of adequate amounts of nucleotide bases dATP, dGTP, dCTP and dTTP, or their analogs, and dideoxynucleotide terminators,

    under such conditions that the DNA template is sequenced.

42.   The method according to claim 41, wherein the DNA polymerase is a modified *Bacillus stearothermophilus* DNA polymerase.

43.   The method according to claim 41, wherein the modified DNA polymerase has an amino

66

acid sequence that shares not less than 95%
homology of a DNA polymerase isolated from a
strain of *Bacillus stearothermophilus, Bacillus
caldotenax* or *Bacillus caldolyticus*.

44. The method according to claim 41,
wherein the modified DNA polymerase has the amino
acid sequence SEQ ID NO:4.

45. The method according to claim 41,
wherein the DNA polymerase is encoded by the
nucleotide sequence SEQ ID NO:3.

46. The method according to claim 41,
wherein the DNA polymerase is a modified DNA
polymerase obtained from a mesophilic bacterium.

47. The method according to claim 41,
wherein the DNA polymerase is a thermostable DNA
polymerase having proofreading 3'-5' exonuclease
activity during DNA sequencing of a DNA strand
from a template, such that the DNA polymerase
functions to excise mismatched nucleotides from
the 3' terminus of the DNA strand at a faster rate
than the rate at which the DNA polymerase
functions to remove nucleotides matched correctly
with nucleotides of the template.

48. The method according to claim 47,
wherein the DNA polymerase is a modified *Bacillus
stearothermophilus, Bacillus caldotenax* or
Bacillus caldolyticus DNA polymerase.

49. A method of radioisotope-labeled
sequencing of a DNA strand comprising the steps
of:
     i)   providing four separate tubes, each

67

containing DNA template to be sequenced, a primer,
dATP, dTTP, dCTP and dGTP, wherein one of dATP,
dTTP, dCTP or dGTP is labeled with a radioisotope,
and DNA polymerase according to claim 1,

    wherein each tube contains an appropriate
amount of one of ddATP, ddTTP, ddCTP or ddGTP;

    ii)   allowing DNA sequencing of the DNA
template to be carried out; and

    iii)   determining the sequence of the DNA
template.

    50.   The method according to claim 49,
wherein the DNA polymerase is a modified *Bacillus
stearothermophilus* DNA polymerase.

    51.   The method according to claim 49,
wherein the modified DNA polymerase has an amino
acid sequence that shares not less than 95%
homology of a DNA polymerase isolated from a
strain of *Bacillus stearothermophilus*, *Bacillus
caldotenax* or *Bacillus caldolyticus*.

    52.   The method according to claim 49,
wherein the modified DNA polymerase has the amino
acid sequence SEQ ID NO:4.

    53.   The method according to claim 49,
wherein the DNA polymerase is encoded by the
nucleotide sequence SEQ ID NO:3.

    54.   The method according to claim 49,
wherein the DNA polymerase is a modified DNA
polymerase obtained from a mesophilic bacterium.

    55.   The method according to claim 49,
wherein the DNA polymerase is a thermostable DNA
polymerase having proofreading 3'-5' exonuclease

activity during DNA sequencing of a DNA strand
from a template, such that the DNA polymerase
functions to excise mismatched nucleotides from
the 3' terminus of the DNA strand at a faster rate
than the rate at which the DNA polymerase
functions to remove nucleotides matched correctly
with nucleotides of the template.

56.   The method according to claim 55,
wherein the DNA polymerase is a modified *Bacillus
stearothermophilus, Bacillus caldotenax* or
Bacillus caldolyticus DNA polymerase.

69
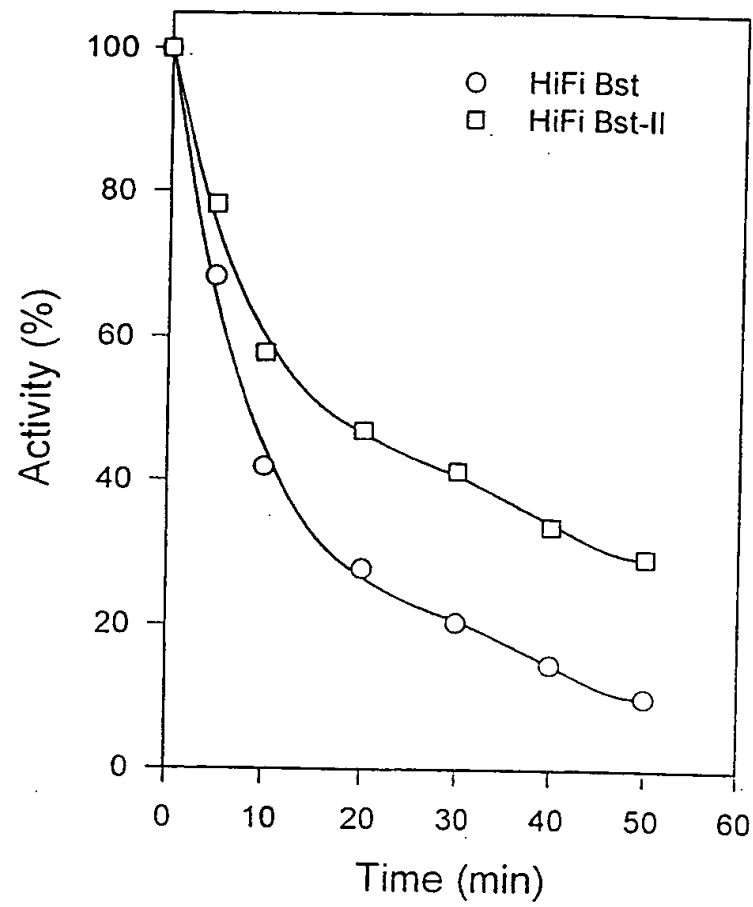
Figure 1

Figure 2

HiFi Bst-II   HiFi Bst

TGCA   TGCA
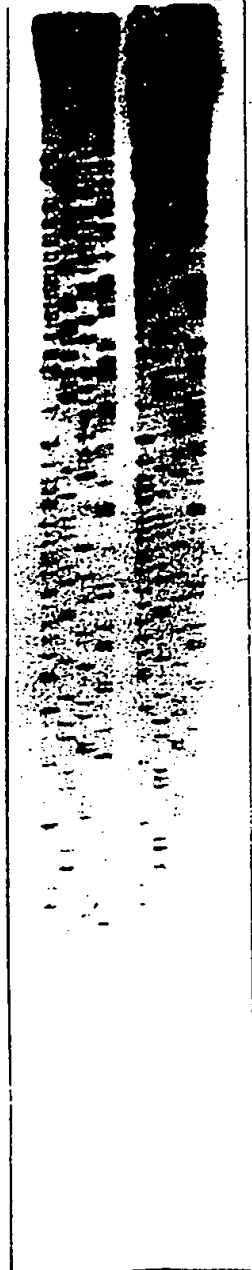
Figure 3
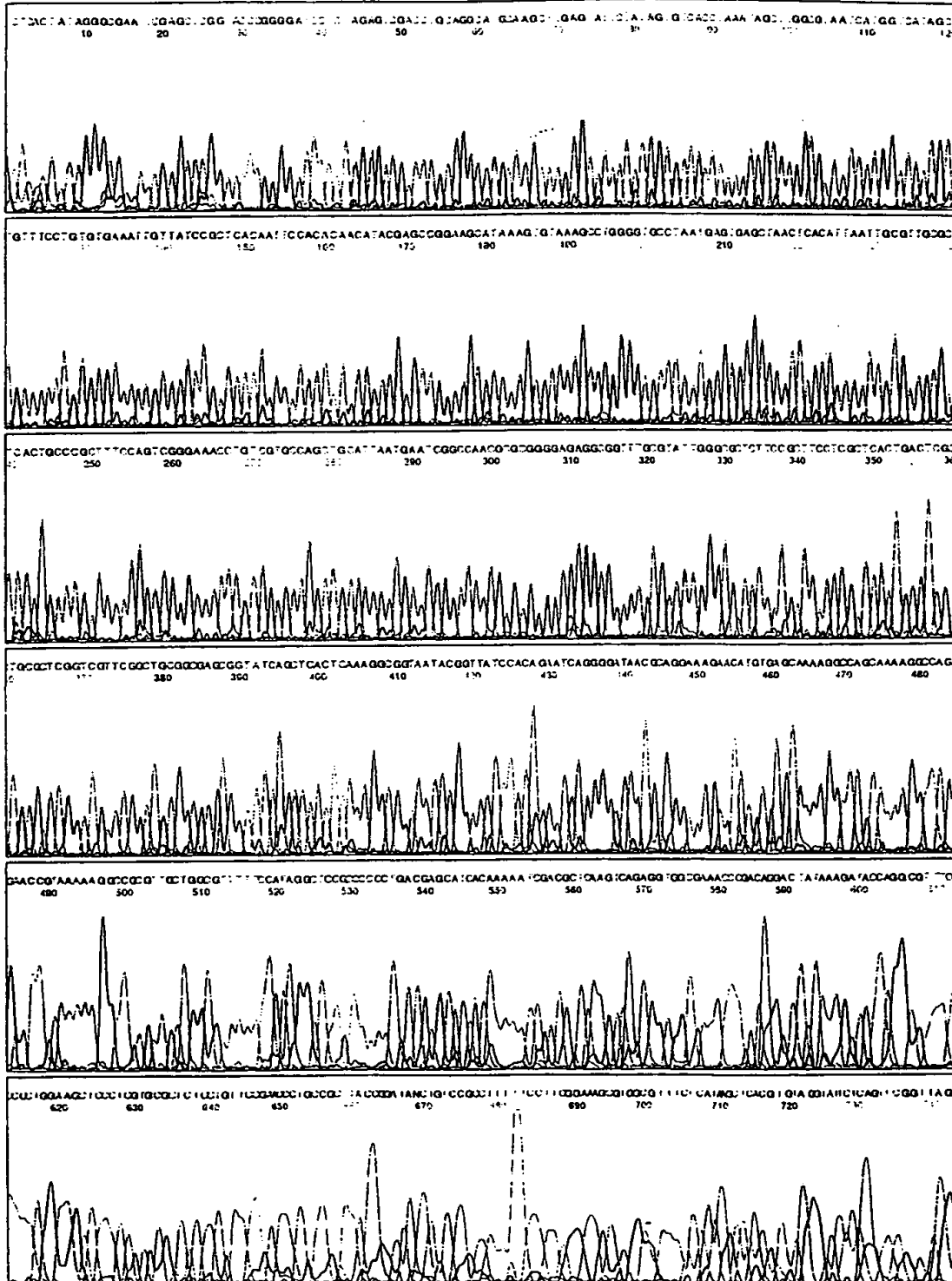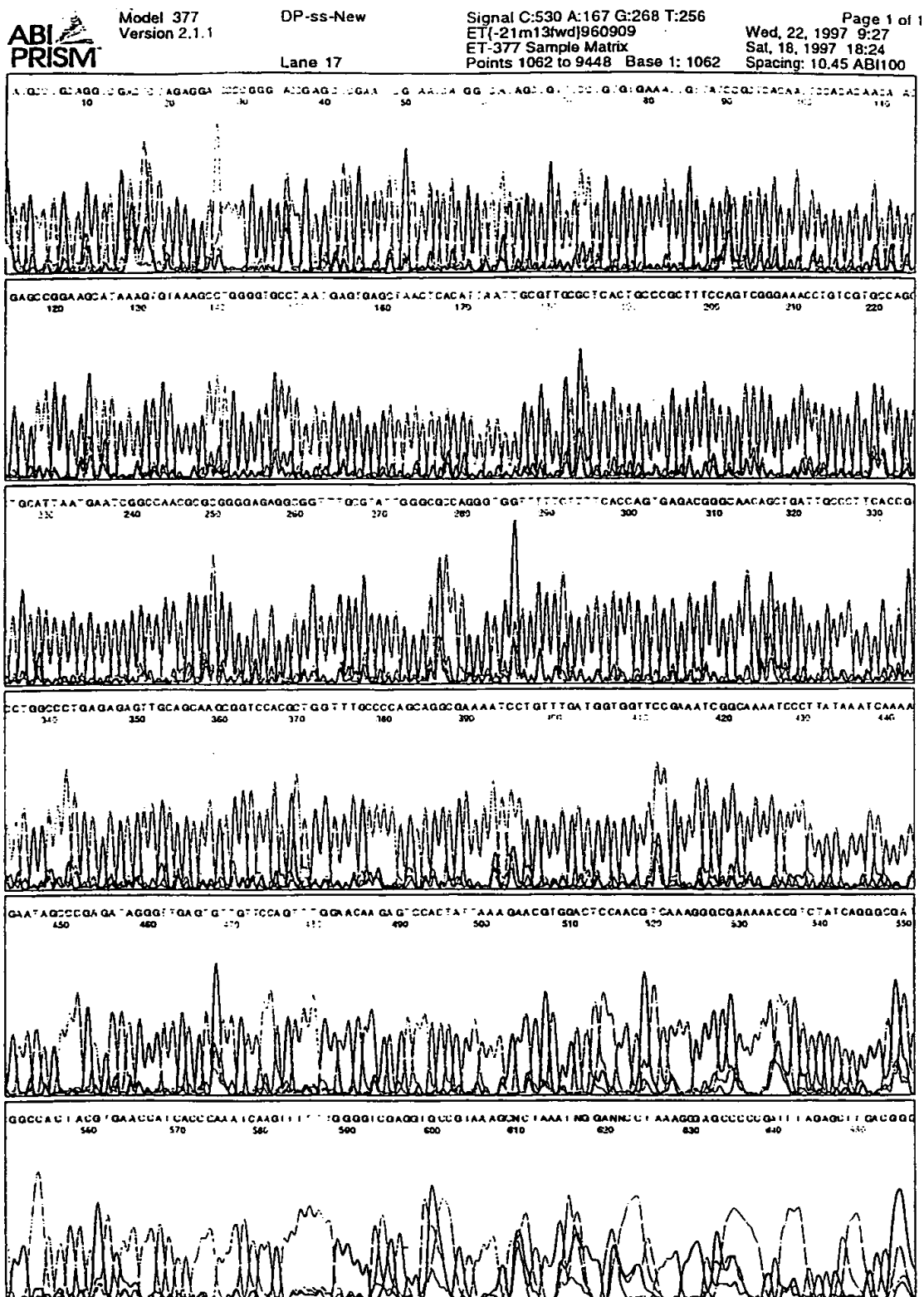
Figure 4

Figure 5
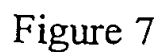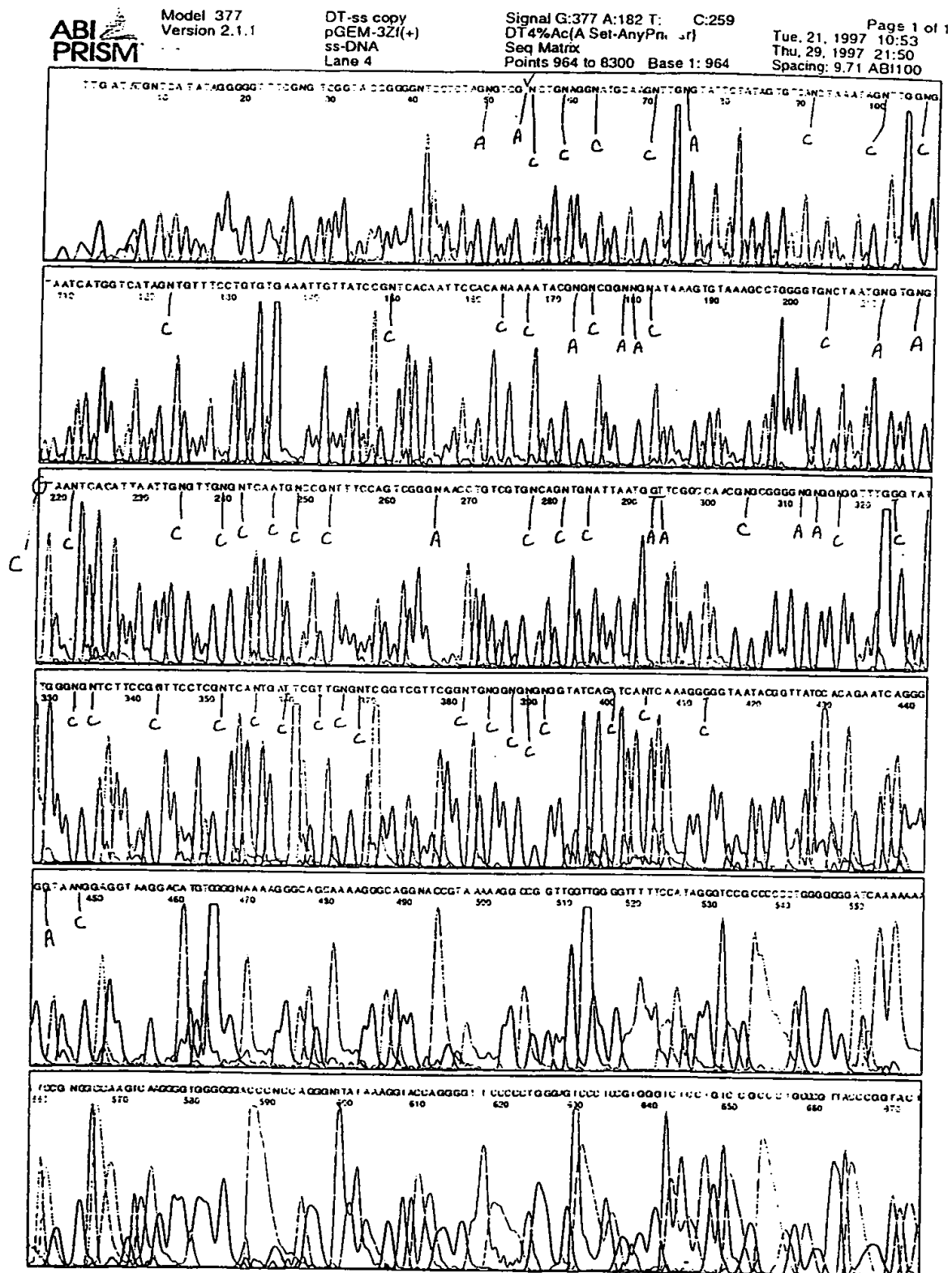
Figure 6

Figure 7

Figure 8

SEQUENCE LISTING

<110> HONG, GUOFAN

HUANG, WEI-HUA

<120> DNA POLYMERASE HAVING ABILITY TO REDUCE INNATE

SELECTIVE DISCRIMINATION AGAINST FLUORESCENT

DYE-LABELED DIDEOXYNUCLEOTIDES

<130> hongsequencelisting

<140> 09/157,397

<141> 1998-09-21

1

<150> 08/544,643

<151> 1995-10-18

<150> 08/642,684

<151> 1996-05-03

<160> 11

<170> PatentIn Ver. 2.0 - beta

<210> 1

<211> 1764

<212> DNA

<213> Bacillus stearothermophilus

2

SUBSTITUTE SHEET (RULE 26)

<400> 1

gccgaagggg agaaaccgct tgaggagatg gagtttgcca tcgttgacgt cattaccgaa 60

gagatgcttg ccgacaaggc agcgcttgtc gttgaggtga tggaagaaaa ctaccacgat 120

gccccgattg tcggaatcgc actagtgaac gagcatgggc gatttttat gcgcccggag 180

accgcgctgg ctgattcgca attttttagca tggcttgccg atgaaacgaa gaaaaaaagc 240

atgtttgacg ccaagcgggc agtcgttgcc ttaaagtgga aaggaattga gcttcgcggc 300

gtcgccttttg atttattgct cgctgcctat ttgctcaatc cggctcaaga tgccggcgat 360

atcgctgcgg tggcgaaaat gaaacaatat gaagcggtgc ggtcggatga agcggtctat 420

ggcaaaggcg tcaagcggtc gctgccggac aacagacgc ttgctgagca tctcgttcgc 480

aaagcggcag ccatttgggc gcttgagcag ccgtttatgg acgatttgcg gaacaacgaa 540

caagatcaat tattaacgaa gcttgagcac gcgctggcgg cgattttggc tgaaatggaa 600

ttcactgggg tgaacgtgga tacaaagcgg cttgaacaga tgggttcgga gctcgccgaa 660

caactgcgtg ccatcgagca gcgcatttac gagctagccg gccaagagtt caacattaac 720

3

SUBSTITUTE SHEET (RULE 26)

```
tcaccaaaac agctcggagt cattttattt gaaaagctgc agctaccggt gctgaagaag 780

acgaaaacag gctattcgac ttcggctgat gtgcttgaga agcttgcgcc gcatcatgaa 840

atcgtcgaaa acattttgca ttaccgccag cttggcaaac tgcaatcaac gtatattgaa 900

ggattgttga aagttgtgcg ccctgatacc ggcaaagtgc atacgatgtt caaccaagcg 960

ctgacgcaaa ctgggcggct cagctcggcc gagccgaact tgcaaaacat tccgattcgg 1020

ctcgaagagg ggcggaaaat ccgccaagcg ttcgtcccgt cagagccgga ctggctcatt 1080

ttcgccgccg attactcaca aattgaattg cgcgtcctcg cccatatcgc cgatgacgac 1140

aatctaattg aagcgttcca acgcgatttg gatattcaca caaaaacggc gatggacatt 1200

ttccagttga gcgaagagga agtcacggcc aacatgcgcc gccaggcaaa ggccgttaac 1260

ttcggtatcg tttacggaat tagcgattac ggattggcgc aaaacttgaa cattacgcgc 1320

aaagaagctg ccgaatttat cgaacgttac ttcgccagct ttccgggcgt aaagcagtat 1380

atggaaaaca tagtgcaaga agcgaaacag aaaggatatg tgacaacgct gttgcatcgg 1440

cgccgctatt tgcctgatat tacaagccgc aatttcaacg tccgcagttt tgcagagcgg 1500

acggccatga acacgccaat tcaaggaagc gccgctgaca ttattaaaaa agcgatgatt 1560
```

4

gatttagcgg cacggctgaa agaagagcag cttcaggctc gtcttttgct gcaagtgcat 1620

gacgagctca ttttggaagc gccaaaagag gaaattgagc gattatgtga gcttgttccg 1680

gaagtgatgg agcaggccgt tacgctccgc gtgccgctga aagtcgacta ccattacggc 1740

ccaacatggt atgatgccaa ataa   1764

<210> 2

<211> 587

<212> PRT

<213> Bacillus stearothermophilus

<400> 2

Ala Glu Gly Glu Lys Pro Leu Glu Glu Met Glu Phe Ala Ile Val Asp

1   5   10   15

5

SUBSTITUTE SHEET (RULE 26)

Val Ile Thr Glu Glu Met Leu Ala Asp Lys Ala Ala Leu Val Val Glu

          20   25   30

Val Met Glu Glu Asn Tyr His Asp Ala Pro Ile Val Gly Ile Ala Leu

          35   40   45

Val Asn Glu His Gly Arg Phe Phe Met Arg Pro Glu Thr Ala Leu Ala

          50   55   60

Asp Ser Gln Phe Leu Ala Trp Leu Ala Asp Glu Thr Lys Lys Lys Ser

65   70   75   80

Met Phe Asp Ala Lys Arg Ala Val Val Ala Leu Lys Trp Lys Gly Ile

          85   90   95

6

SUBSTITUTE SHEET (RULE 26)

Glu Leu Arg Gly Val Ala Phe Asp Leu Leu Leu Ala Ala Tyr Leu Leu

         100   105   110

Asn Pro Ala Gln Asp Ala Gly Asp Ile Ala Ala Val Ala Lys Met Lys

         115   120   125

Gln Tyr Glu Ala Val Arg Ser Asp Glu Ala Val Tyr Gly Lys Gly Val

         130   135   140

Lys Arg Ser Leu Pro Asp Glu Gln Thr Leu Ala Glu His Leu Val Arg

145   150   155   160

Lys Ala Ala Ala Ile Trp Ala Leu Glu Gln Pro Phe Met Asp Asp Leu

7

SUBSTITUTE SHEET (RULE 26)

165   170   175

Arg Asn Asn Glu Gln Asp Gln Leu Leu Thr Lys Leu Glu His Ala Leu

            180   185   190

Ala Ala Ile Leu Ala Glu Met Glu Phe Thr Gly Val Asn Val Asp Thr

        195   200   205

Lys Arg Leu Glu Gln Met Gly Ser Glu Leu Ala Glu Gln Leu Arg Ala

        210   215   220

Ile Glu Gln Arg Ile Tyr Glu Leu Ala Gly Gln Glu Phe Asn Ile Asn

225   230   235   240

8

Ser Pro Lys Gln Leu Gly Val Ile Leu Phe Glu Lys Leu Gln Leu Pro

245   250   255

Val Leu Lys Lys Thr Lys Thr Gly Tyr Ser Thr Ser Ala Asp Val Leu

260   265   270

Glu Lys Leu Ala Pro His His Glu Ile Val Glu Asn Ile Leu His Tyr

275   280   285

Arg Gln Leu Gly Lys Leu Gln Ser Thr Tyr Ile Glu Gly Leu Leu Lys

290   295   300

Val Val Arg Pro Asp Thr Gly Lys Val His Thr Met Phe Asn Gln Ala

305   310   315   320

9

Leu Thr Gln Thr Gly Arg Leu Ser Ser Ala Glu Pro Asn Leu Gln Asn

        325   330   335

Ile Pro Ile Arg Leu Glu Glu Gly Arg Lys Ile Arg Gln Ala Phe Val

        340   345   350

Pro Ser Glu Pro Asp Trp Leu Ile Phe Ala Ala Asp Tyr Ser Gln Ile

        355   360   365

Glu Leu Arg Val Leu Ala His Ile Ala Asp Asp Asp Asn Leu Ile Glu

        370   375   380

Ala Phe Gln Arg Asp Leu Asp Ile His Thr Lys Thr Ala Met Asp Ile

                               10

385   390   395   400

Phe Gln Leu Ser Glu Glu Glu Val Thr Ala Asn Met Arg Arg Gln Ala

               405   410   415

Lys Ala Val Asn Phe Gly Ile Val Tyr Gly Ile Ser Asp Tyr Gly Leu

               420   425   430

Ala Gln Asn Leu Asn Ile Thr Arg Lys Glu Ala Ala Glu Phe Ile Glu

               435   440   445

Arg Tyr Phe Ala Ser Phe Pro Gly Val Lys Gln Tyr Met Glu Asn Ile

               450   455   460

11

Val Gln Glu Ala Lys Gln Lys Gly Tyr Val Thr Thr Leu Leu His Arg

465   470   475   480

Arg Arg Tyr Leu Pro Asp Ile Thr Ser Arg Asn Phe Asn Val Arg Ser

485   490   495

Phe Ala Glu Arg Thr Ala Met Asn Thr Pro Ile Gln Gly Ser Ala Ala

500   505   510

Asp Ile Ile Lys Lys Ala Met Ile Asp Leu Ala Ala Arg Leu Lys Glu

515   520   525

Glu Gln Leu Gln Ala Arg Leu Leu Leu Gln Val His Asp Glu Leu Ile

530   535   540

12

Leu Glu Ala Pro Lys Glu Glu Ile Glu Arg Leu Cys Glu Leu Val Pro

545   550   555   560

Glu Val Met Glu Gln Ala Val Thr Leu Arg Val Pro Leu Lys Val Asp

                    565   570   575

Tyr His Tyr Gly Pro Thr Trp Tyr Asp Ala Lys

          580   585

<210> 3

<211> 1764

<212> DNA

<213> Bacillus stearothermophilus

<400> 3

atggccgaag gggagaaacc gcttgaggag atggagtttg ccatcgttga cgtcattacc 60

gaagagatgc ttgccgacaa ggcagcgctt gtcgttgagg tgatggaaga aaactaccac 120

gatgccccga ttgtcggaat cgcactagtg aacgagcatg gcgattttt tatgcgcccg 180

gagaccgcgc tggctgattc gcaattttta gcatggcttg ccgatgaaac gaagaaaaaa 240

agcatgtttg acgccaagcg ggcagtcgtt gccttaaagt ggaaaggaat tgagcttcgc 300

ggcgtcgcct ttgatttatt gctcgctgcc tatttgctca atccggctca agatgccggc 360

gatatcgctg cggtggcgaa aatgaaacaa tatgaagcgg tgcggtcgga tgaagcggtc 420

tatggcaaag cgtcaagcg gtcgctgccg acgaacaga cgcttgctga gcatctcgtt 480

cgcaaagcgg cagccatttg ggcgcttgag cagccgttta tggacgattt gcggaacaac 540

gaacaagatc aattattaac gaagcttgag cacgcgctgg cggcgatttt ggctgaaatg 600

gaattcactg gggtgaacgt ggatacaaag cggcttgaac agatgggttc ggagctcgcc 660

14

gaacaactgc gtgccatcga gcagcgcatt tacgagctag ccggccaaga gttcaacatt 720

aactcaccaa aacagctcgg agtcatttta tttgaaaagc tgcagctacc ggtgctgaag 780

aagacgaaaa caggctattc gacttcggct gatgtgcttg agaagcttgc gccgcatcat 840

gaaatcgtcg aaaacatttt gcattaccgc cagcttggca aactgcaatc aacgtatatt 900

gaaggattgt tgaaagttgt gcgccctgat accggcaaag tgcatacgat gttcaaccaa 960

gcgctgacgc aaactgggcg gctcagctcg gccgagccga acttgcaaaa cattccgatt 1020

cggaccccac tggggcggaa aatccgccaa gcgttcgtcc cgtcagagcc ggactggctc 1080

attttcgccg ccgattactc acaaattgaa ttgcgcgtcc tcgcccatat cgccgatgac 1140

gacaatctaa ttgaagcgtt ccaacgcgat ttggatattc acacaaaaac ggcgatggac 1200

attttccagt tgagcgaaga ggaagtcacg gccaacatgc gccgccaggc aaaggccgtt 1260

aactacggta tcgtttacgg aattagcgat tacggattgg cgcaaaactt gaacattacg 1320

cgcaaagaag ctgccgaatt tatcgaacgt tacttcgcca gctttccggg cgtaaagcag 1380

tatatggaaa acatagtgca agaagcgaaa cagaaaggat atgtgacaac gctgttgcat 1440

cggcgccgct atttgcctga tattacaagc cgcaatttca cgtccgcag ttttgcagag 1500

15

cggacggcca tgaacacgcc aattcaagga agcgccgctg acattattaa aaaagcgatg 1560

attgatttag cggcacggct gaaagaagag cagcttcagg ctcgtctttt gctgcaagtg 1620

catgacgagc tcattttgga agcgccaaaa gaggaaattg agcgattatg tgagcttgtt 1680

ccggaagtga tggagcaggc cgttacgctc cgcgtgccgc tgaaagtcga ctaccattac 1740

ggcccaacat ggtatgatgc caaa   1764

<210> 4

<211> 588

<212> PRT

<213> Bacillus stearothermophilus

<400> 4

Met Ala Glu Gly Glu Lys Pro Leu Glu Glu Met Glu Phe Ala Ile Val

1   5   10   15

16

Asp Val Ile Thr Glu Glu Met Leu Ala Asp Lys Ala Ala Leu Val Val

20   25   30

Glu Val Met Glu Glu Asn Tyr His Asp Ala Pro Ile Val Gly Ile Ala

35   40   45

Leu Val Asn Glu His Gly Arg Phe Phe Met Arg Pro Glu Thr Ala Leu

50   55   60

Ala Asp Ser Gln Phe Leu Ala Trp Leu Ala Asp Glu Thr Lys Lys Lys

65   70   75   80

Ser Met Phe Asp Ala Lys Arg Ala Val Val Ala Leu Lys Trp Lys Gly

17

85   90   95

Ile Glu Leu Arg Gly Val Ala Phe Asp Leu Leu Leu Ala Ala Tyr Leu

          100   105   110

Leu Asn Pro Ala Gln Asp Ala Gly Asp Ile Ala Ala Val Ala Lys Met

          115   120   125

Lys Gln Tyr Glu Ala Val Arg Ser Asp Glu Ala Val Tyr Gly Lys Gly

          130   135   140

Val Lys Arg Ser Leu Pro Asp Glu Gln Thr Leu Ala Glu His Leu Val

145   150   155   160

18

Arg Lys Ala Ala Ala Ile Trp Ala Leu Glu Gln Pro Phe Met Asp Asp

165   170   175

Leu Arg Asn Asn Glu Gln Asp Gln Leu Leu Thr Lys Leu Glu His Ala

180   185   190

Leu Ala Ala Ile Leu Ala Glu Met Glu Phe Thr Gly Val Asn Val Asp

195   200   205

Thr Lys Arg Leu Glu Gln Met Gly Ser Glu Leu Ala Glu Gln Leu Arg

210   215   220

Ala Ile Glu Gln Arg Ile Tyr Glu Leu Ala Gly Gln Glu Phe Asn Ile

225   230   235   240

Asn Ser Pro Lys Gln Leu Gly Val Ile Leu Phe Glu Lys Leu Gln Leu

                    245   250   255

Pro Val Leu Lys Lys Thr Lys Thr Gly Tyr Ser Thr Ser Ala Asp Val

                    260   265   270

Leu Glu Lys Leu Ala Pro His His Glu Ile Val Glu Asn Ile Leu His

                    275   280   285

Tyr Arg Gln Leu Gly Lys Leu Gln Ser Thr Tyr Ile Glu Gly Leu Leu

                    290   295   300

Lys Val Val Arg Pro Asp Thr Gly Lys Val His Thr Met Phe Asn Gln

                                        20

305   310   315   320

Ala Leu Thr Gln Thr Gly Arg Leu Ser Ser Ala Glu Pro Asn Leu Gln

         325   330   335

Asn Ile Pro Ile Arg Thr Pro Leu Gly Arg Lys Ile Arg Gln Ala Phe

         340   345   350

Val Pro Ser Glu Pro Asp Trp Leu Ile Phe Ala Ala Asp Tyr Ser Gln

         355   360   365

Ile Glu Leu Arg Val Leu Ala His Ile Ala Asp Asp Asp Asn Leu Ile

         370   375   380

21

Glu Ala Phe Gln Arg Asp Leu Asp Ile His Thr Lys Thr Ala Met Asp

385   390   395   400

Ile Phe Gln Leu Ser Glu Glu Glu Val Thr Ala Asn Met Arg Arg Gln

                  405   410   415

Ala Lys Ala Val Asn Tyr Gly Ile Val Tyr Gly Ile Ser Asp Tyr Gly

            420   425   430

Leu Ala Gln Asn Leu Asn Ile Thr Arg Lys Glu Ala Ala Glu Phe Ile

            435   440   445

Glu Arg Tyr Phe Ala Ser Phe Pro Gly Val Lys Gln Tyr Met Glu Asn

            450   455   460

22

Ile Val Gln Glu Ala Lys Gln Lys Gly Tyr Val Thr Thr Leu Leu His

465   470   475   480

Arg Arg Arg Tyr Leu Pro Asp Ile Thr Ser Arg Asn Phe Asn Val Arg

                485   490   495

Ser Phe Ala Glu Arg Thr Ala Met Asn Thr Pro Ile Gln Gly Ser Ala

                500   505   510

Ala Asp Ile Ile Lys Lys Ala Met Ile Asp Leu Ala Ala Arg Leu Lys

                515   520   525

Glu Glu Gln Leu Gln Ala Arg Leu Leu Leu Gln Val His Asp Glu Leu

                530   535   540

Ile Leu Glu Ala Pro Lys Glu Glu Ile Glu Arg Leu Cys Glu Leu Val

545   550   555   560

Pro Glu Val Met Glu Gln Ala Val Thr Leu Arg Val Pro Leu Lys Val

                          565   570   575

Asp Tyr His Tyr Gly Pro Thr Trp Tyr Asp Ala Lys

                580   585

<210> 5

<211> 17

<212> DNA

<213> Bacillus stearothermophilus

<400> 5

cattttgctg ccggtca   17

<210> 6

<211> 20

<212> DNA

<213> Bacillus stearothermophilus

<400> 6

gtaaaacgac ggccagtctt   20

<210> 7

<211> 20

<212> DNA

<213> Bacillus stearothermophilus

<400> 7

gtaaaacgac ggccagtcgg   20

<210> 8

<211> 20

<212> DNA

<213> Bacillus stearothermophilus

<400> 8

cattttgctg ccggtcagaa   20

<210> 9

<211> 26

<212> DNA

<213> Bacillus stearothermophilus

<400> 9

gccgttaact acggtatcgt ttacgg   26

<210> 10

<211> 36

<212> DNA

<213> Bacillus stearothermophilus

<400> 10

cattccgatt cggaccccac tgggccggaa aatccg   36

<210> 11

<211> 24

<212> DNA

<213> Bacillus stearothermophilus

<400> 11

ssssssstgta aaacgacggc cagt   24

28

A.	CLASSIFICATION OF SUBJECT MATTER

IPC⁶ :C12N9/22, C07K14/195,C12N15/55,C12Q1/68

According to International Patent Classification(IPC) or to both national classification and IPC

B.	FIELDS SEARCHED

Minimum documentation searched(classification system followed by classification symbols)

Documentation searched other than minimum documentation to the extent that such documents are included in the field searched

Electronic data base consulted during the international search(name of data base and, where practicable, search terms used)
GenBank, EMBL, DDBJ, PDB, SwissProt, SPupdate, PIR

C.	DOCUMENTS CONSIDERED TO BE RELEVANT	
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant claim No.
A	GenBank U33536 21.APR.1997 see the sequence. or in: Genet Anal 1996 Mar;12(5-6):185-95 Aliotta JM,et al: " Thermostable Bst DNA polymerase I lacks a 3'-->5' proofreading exonuclease activity." see the abstract.	1-56
A	GenBank U93028 02.JUL.1997 see the sequence.	1-56

☒ Further documents are listed in the continuation of Box C.   ☐   See patent family annex.

* Special categories of cited documents:	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A" document defining the general state of the art which is not considered to be of particular relevance	
"E" earlier document but published on or after the international filing date	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason(as specified)	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"O" document referring to an oral disclosure, use, exhibition or other means	
"P" document published prior to the international filing date but later than the priority date claimed	"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
21.APR.1999 (21.04.99)	1 7 JUN 1999 (1 7. 06. 9 9)

Name and mailing address of the ISA/	
The Chinese Patent Office 6, Xitucheng Road, Haidian District, Beijing, 100088, China	Authorized officer ZENG,Fanhui
	Telephone No. 86-010-62093733
Facsimile No. 86-010-62019451	

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	Biochim Biophys Acta 1996 Jun 7;1307(2):178-86 Riggs MG,et al." Construction of single amino acid substitution mutants of cloned Bacillus stearothermophilus DNA polymerase I which lack 5'-->3' exonuclease activity." see the abstract.	1-56
A	J Biochem (Tokyo) 1993 Mar;113(3):401-410 Uemori T,et al:" Cloning of the DNA polymerase gene of Bacillus caldotenax and characterization of the gene product." see the abstract.	1-56
A	GenBank U23149 24.JAN.1996 see the sequence. or in: Gene 1995 Sep 22;163(1):65-8 Phang SM,et al: " Cloning and complete sequence of the DNA polymerase-encoding gene (Bstpoll) and characterisation of the Klenow-like fragment from Bacillus stearothermophilus." see the abstract.	1-56